High resolution imaging with LOFAR

Neal Jackson

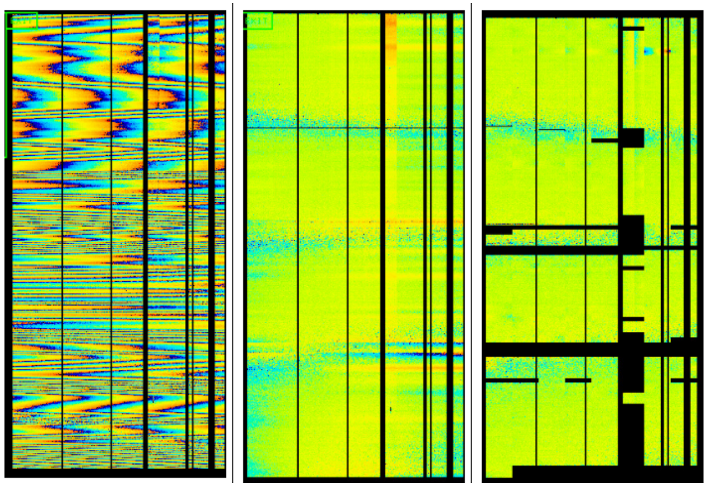
(with thanks to the LOFAR Long-Baseline Working Group)

1. VLBI imaging problems
2. The LOFAR long baselines: science and surveys
3. LBCS: the LOFAR Long-Baseline Calibrator Survey
4. LOFAR-LB imaging: towards a pipeline

1. VLBI imaging problems

a) Why it is more difficult

Effects on the data



Phase against
Time (y-axis)
Frequency (x-axis)

DELAYS

- give phase variation with frequency

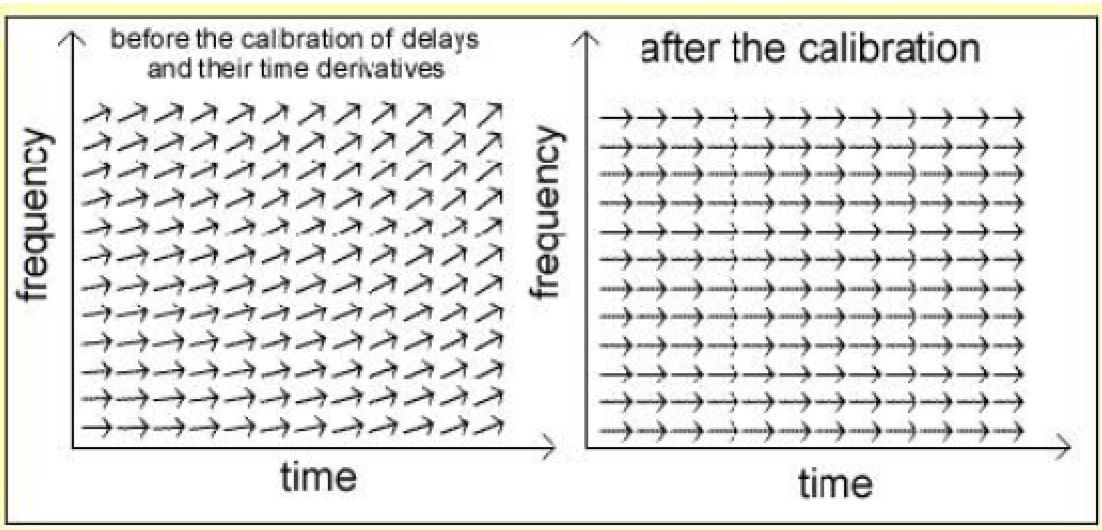
Non-dispersive: clock errors
Constant time vs. frequency
d(phi)/df constant

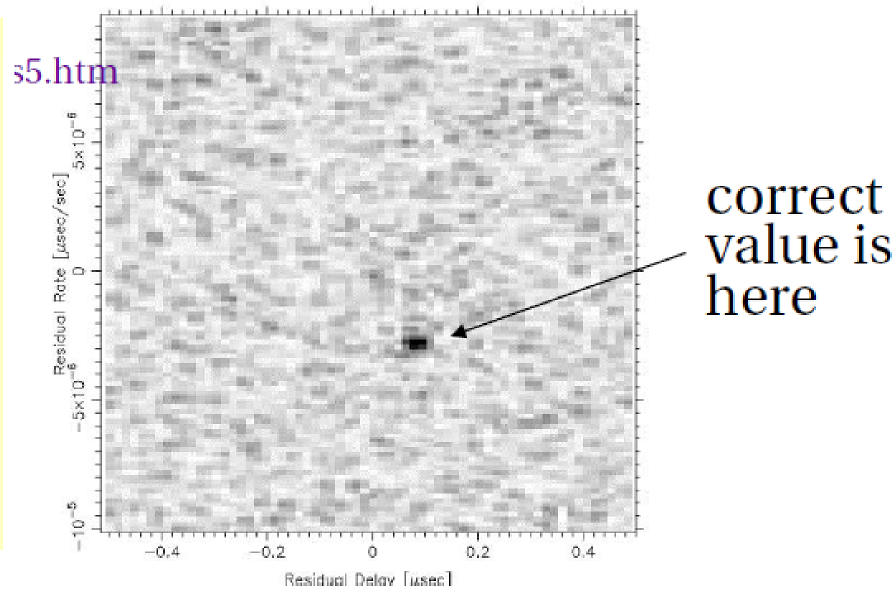
Dispersive: ionosphere
Effective delay increases with decreasing f

PHASE

- phase variation with time

Rapid changes in atmosphere (ionosphere at LF)





Aim is to take out both effects on phase

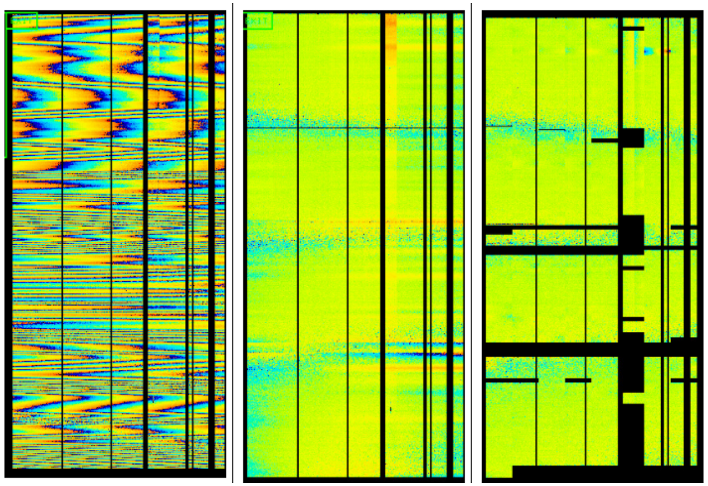
Remnant phase structure is a (distorted) FT of the source – more of this later

Note that phase can vary rapidly with time!

Table 4.5: Coherence times on the longest baseline of MERLIN

Frequency	Good	Bad
151 MHz	3 min.	15 s
408 MHz	10 min	45 s
1.4/1.7 GHz	40 min.	1 min.
5 GHz	40 min	5 min
22 GHz	10 min	30 s

Effects on the data



Phase against
  Time (y-axis)
  Frequency (x-axis)

Corrected data in centre/right

CORRECTING THESE EFFECTS

"Fringe-fitting" and phase calibration

Aim is to determine, for each antenna:
- delay as function of time/frequency
- phase

Need a calibrator
- close to source for phase and dispersive delay
- can be further away for clock correction

IF YOU DON'T CORRECT...

- delays → cannot average over frequency
- phases → cannot average over time

NEED CALIBRATOR SOURCE(s)

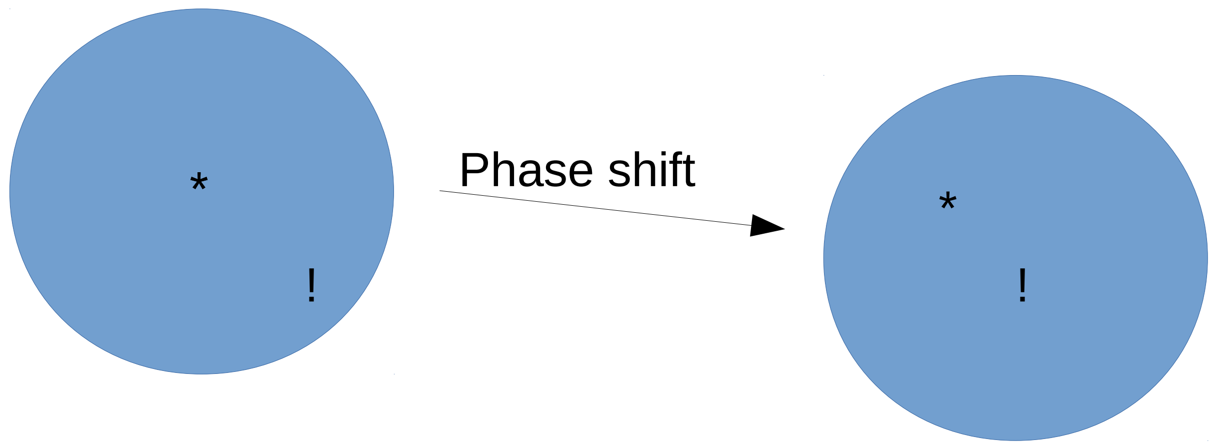
- close to target
- with structure on correct scales
- motivation for LBCS (later)

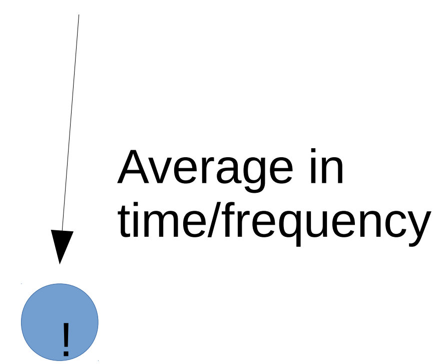
1. VLBI imaging problems

b) Why it is easier

With long baselines, can use bandwidth and time smearing to your advantage:

Start with wide-field dataset





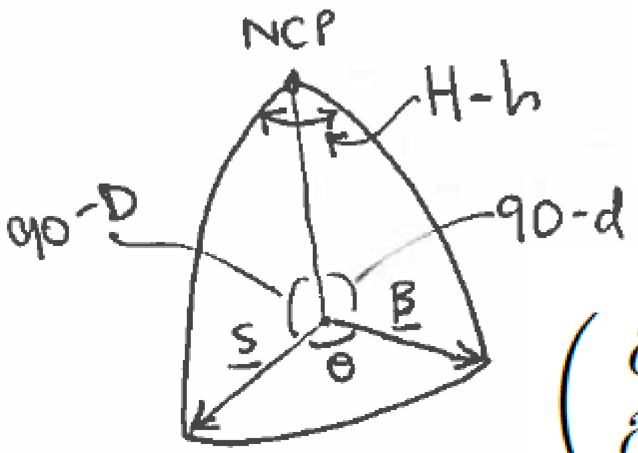
Phase shift dataset to a different position
Then average in time/frequency

2 advantages:
- dataset becomes much smaller
- can isolate sources; here, can map
  source ! without interference from source *

But need to be very careful when averaging – only average over times and frequencies which are coherent! (So if not calibrated = atm. Coherence time / delay length)

Fringe-rate/delay mapping

With atmospheric/clock effects removed, phase as a function of time/frequency allows a map with a single baseline by Fourier transform



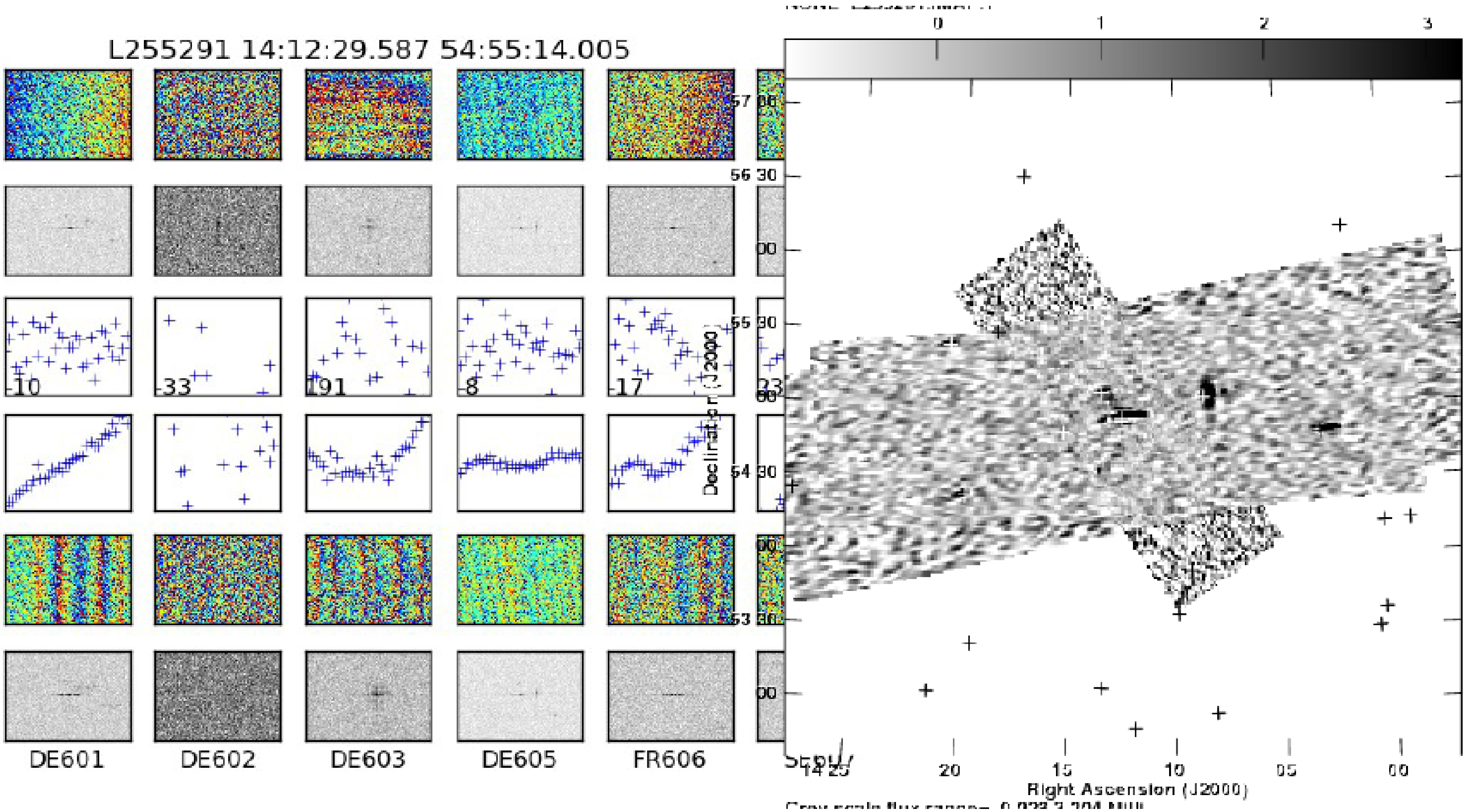
$$\phi = \frac{2\pi L}{\lambda}\left(\sin d \sin D + \cos d \cos D \cos(H-h)\right)$$

$$\begin{pmatrix} \partial\phi/\partial t \\ \partial\phi/\partial f \end{pmatrix} = \begin{pmatrix} \frac{1}{\cos D}\frac{\partial^2\phi}{\partial H \partial t} & \frac{\partial^2\phi}{\partial D \partial t} \\ \frac{1}{\cos D}\frac{\partial^2\phi}{\partial H \partial f} & \frac{\partial^2\phi}{\partial D \partial f} \end{pmatrix} \begin{pmatrix} \Delta H \\ \Delta D \end{pmatrix}$$

Phase slope in t and f related to structure, by a transformation matrix
This matrix includes the second derivatives of the interferometer phase

Fringe rate / delay mapping in practice

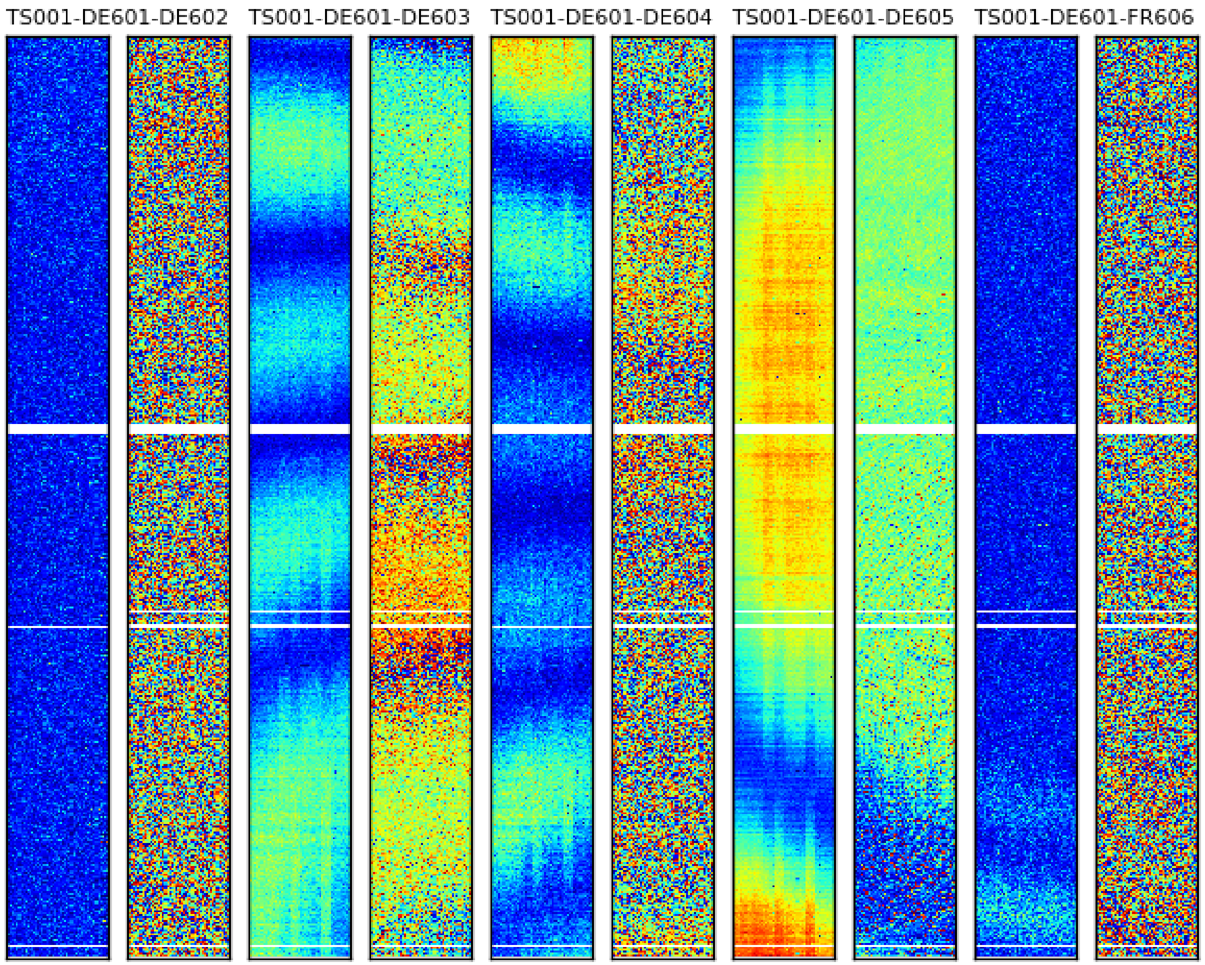
LOFAR LBCS – baselines to ST001



Resolution in time direction depends on time window (and hence time coherence)
Resolution in frequency direction depends on frequency window (and delay coherence)

1. VLBI imaging problems

c) Final remark: closure phase



UNCALIBRATED DATA!

But with phase 1-2 added to 2-3 added to 3-1

Phase variation effects cancel

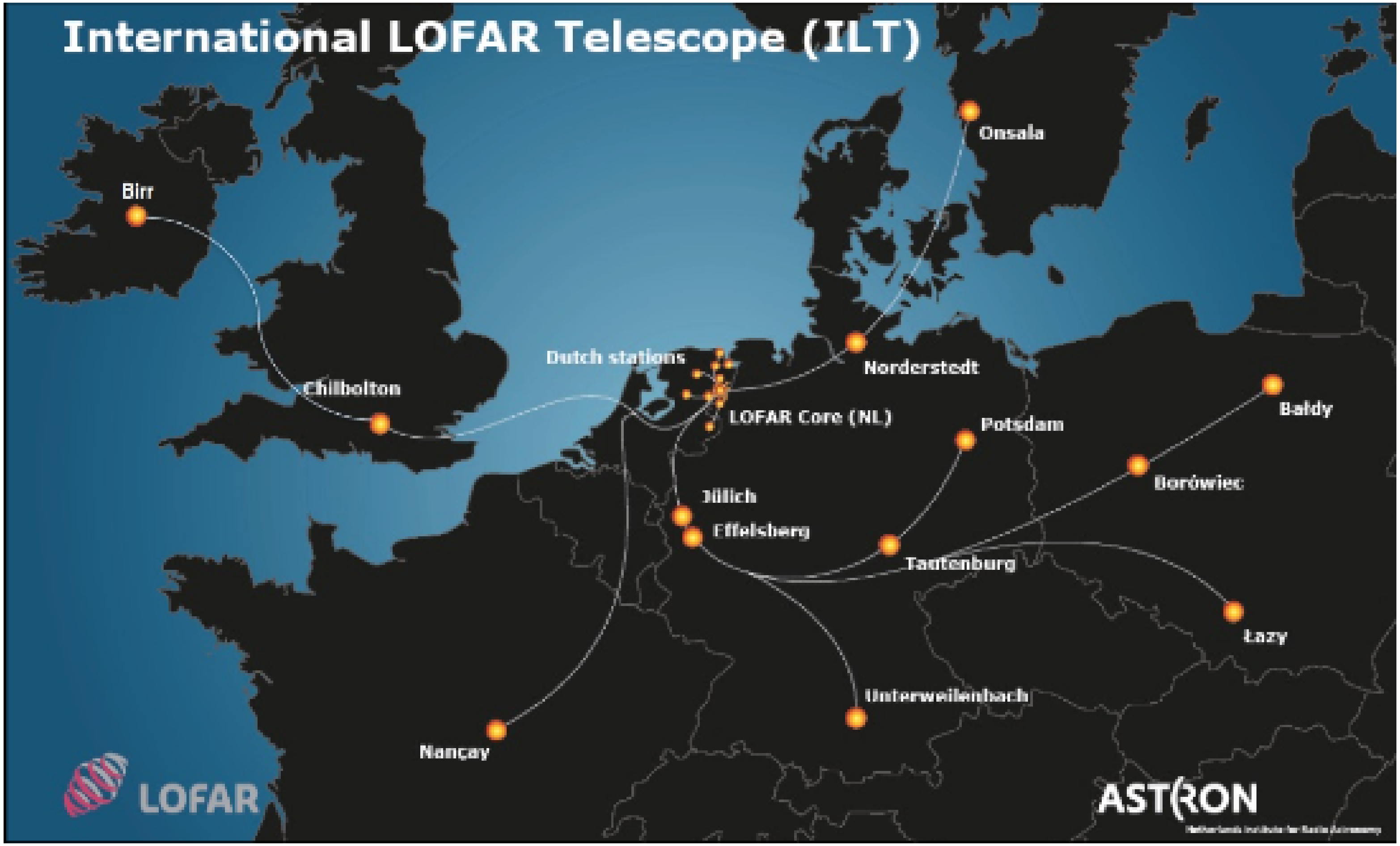
Leaves only source information

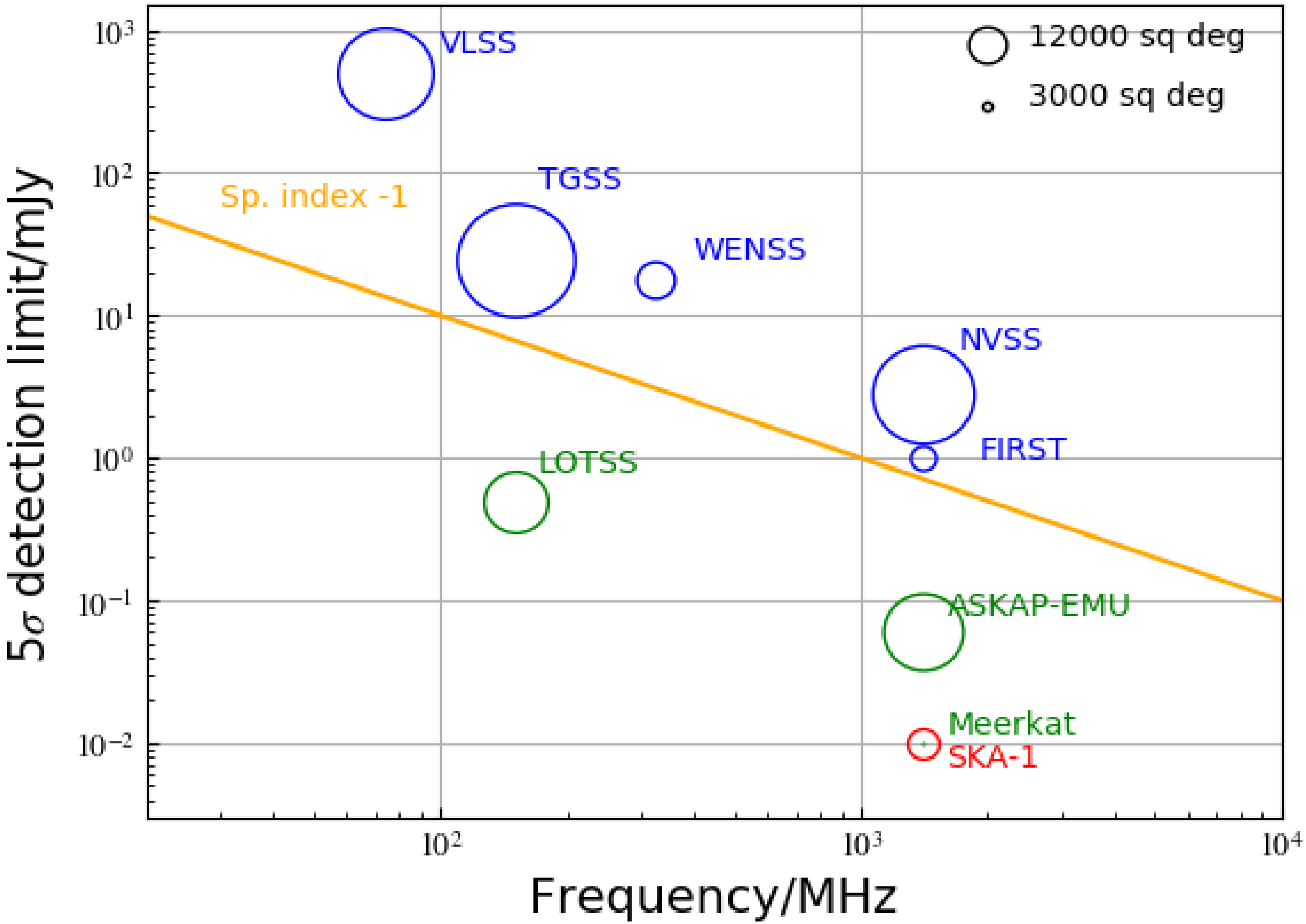
(Could map using only CP, but this is very hard in practice)

The University of Manchester
Jodrell Bank
Observatory

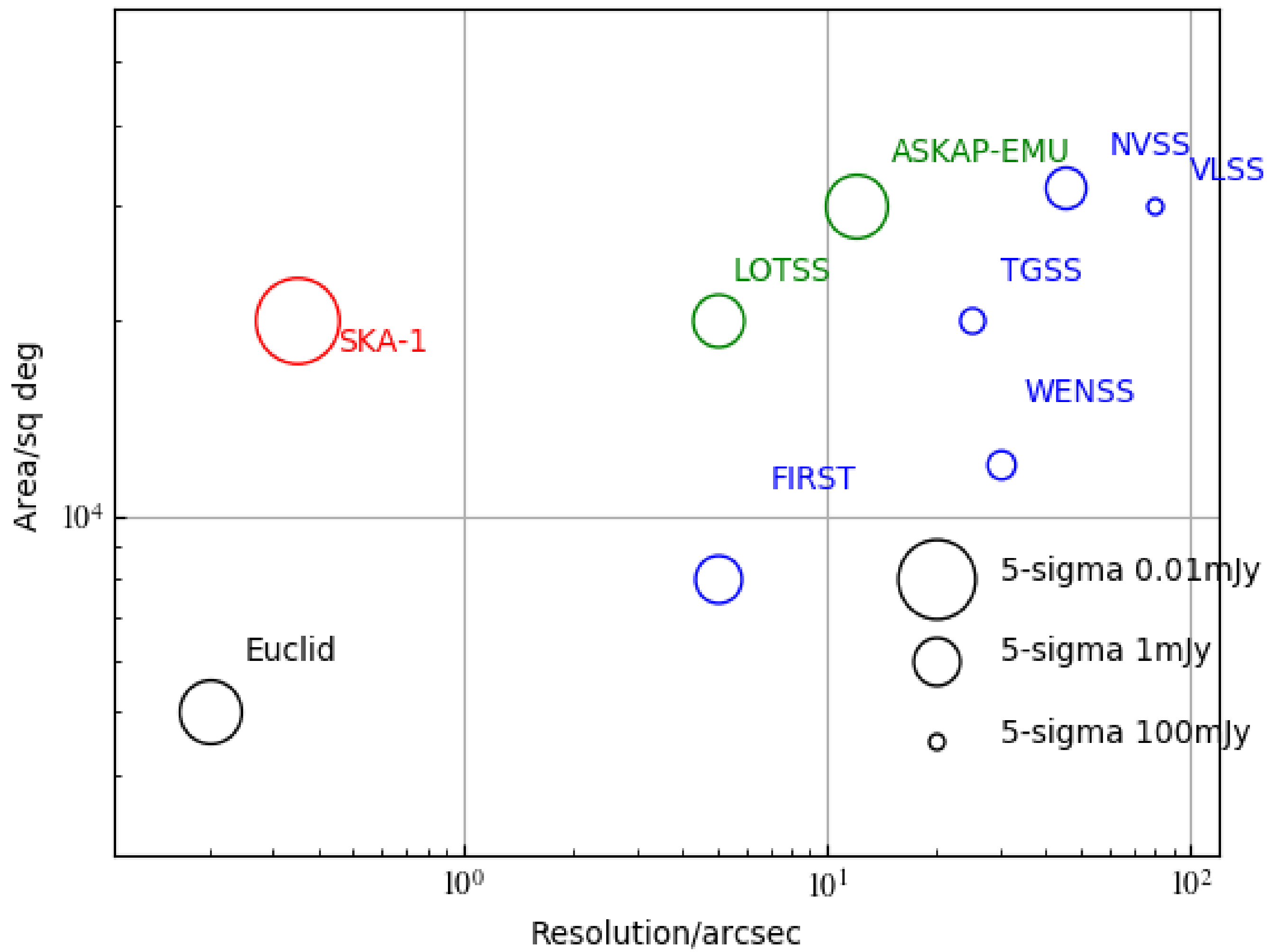
MANCHESTER
1824

2. LOFAR long baselines: science and surveys
(see talks by Martin Hardcastle, Leah Morabito, Rachael Ainsworth)

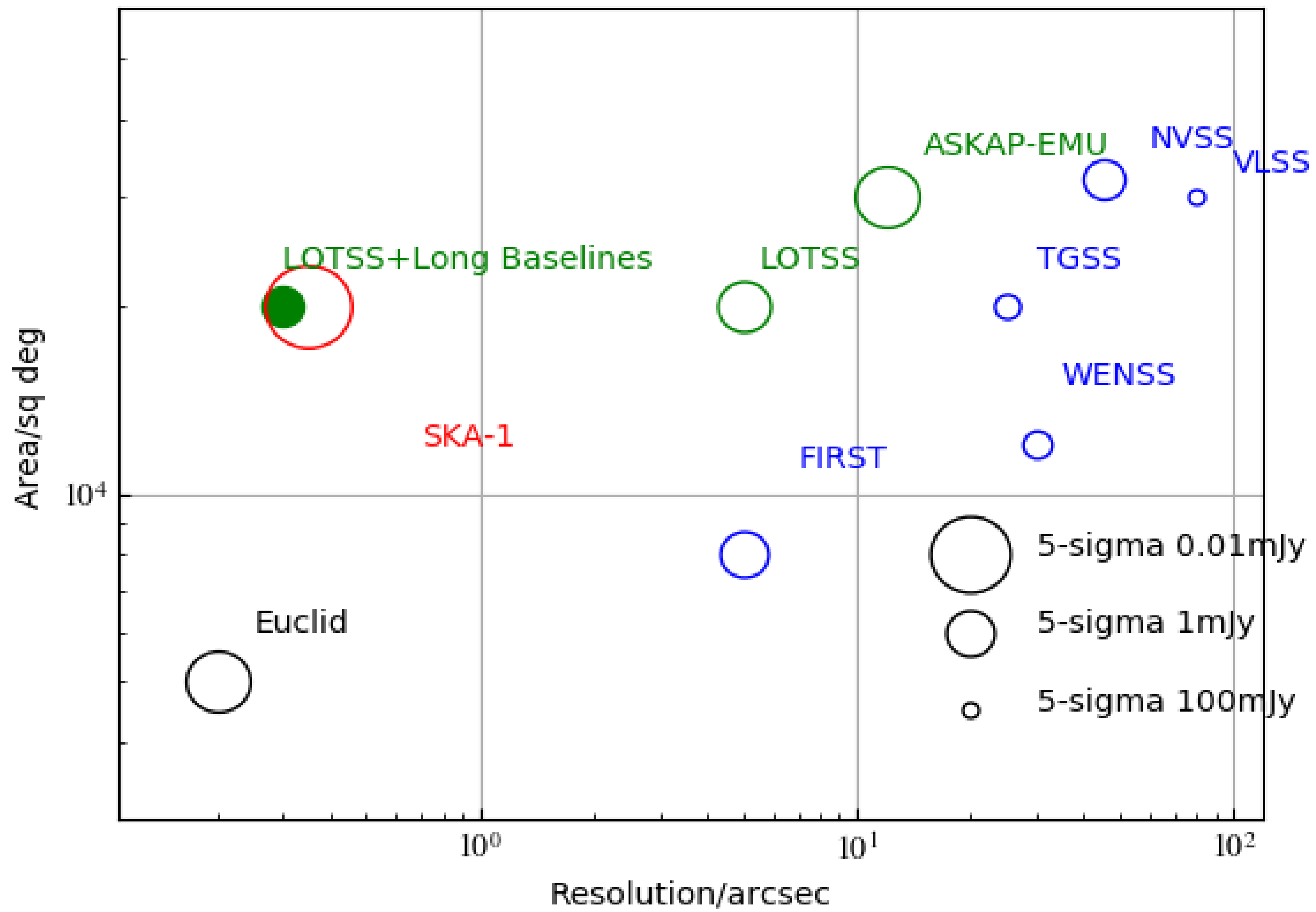




Conventional survey space: depth vs frequency
Traditional 'wedding-cake' structure: sacrifice area for depth
Current, in-progress and future surveys separated by factor ~10 in depth
Combination of depth and multiple frequencies gives science on individual objects



Alternative 'survey pixel volume' space
Higher resolution makes huge areas very difficult (even for SKA)
Relevant survey space for e.g. gravitational lens surveying, parallel studies of many objects



Alternative 'survey pixel volume' space
Higher resolution makes huge areas very difficult (even for SKA)
Relevant survey space for e.g. gravitational lens surveying, parallel studies of many objects
Extension of LOTSS to long baselines can be a leading survey in this space (even with SKA1)

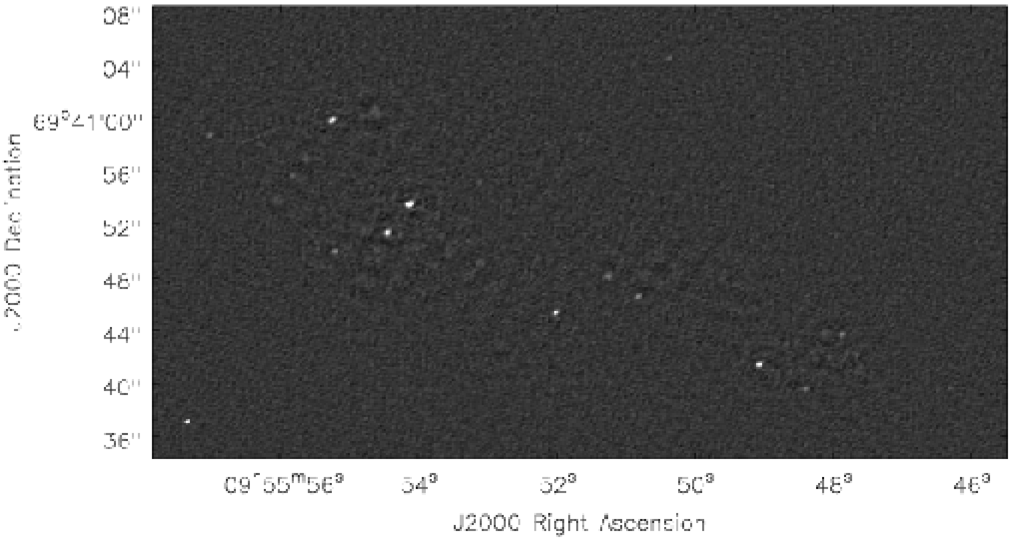
LOFAR LB survey gets you a wide-area survey of sources

At sub-galactic resolution for most z

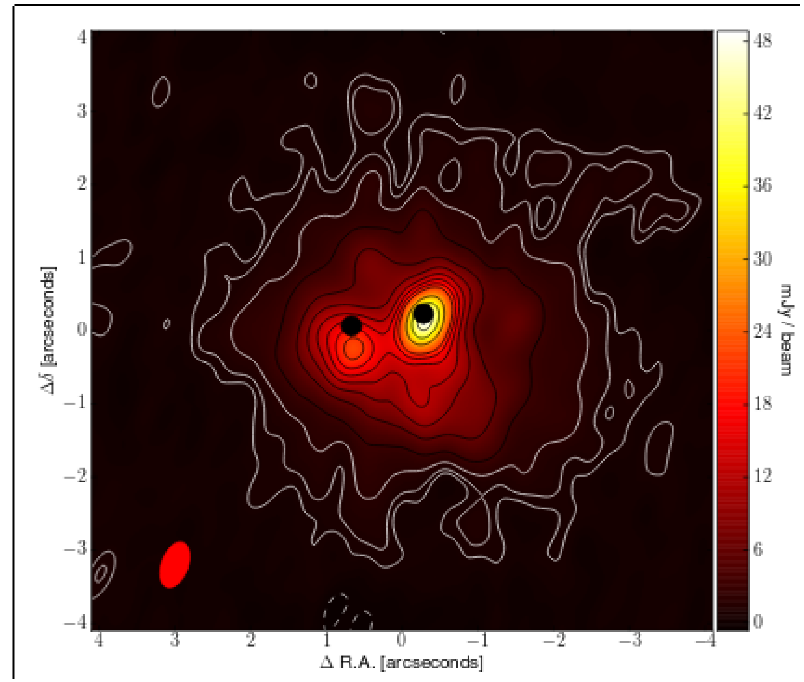
To highly competitive (though <SKA-1) sensitivity for steep-spectrum sources (but with complementary science, and earlier)

And 1-1.2 times resolution of SKA-1

Science with the long baselines: star-forming galaxies



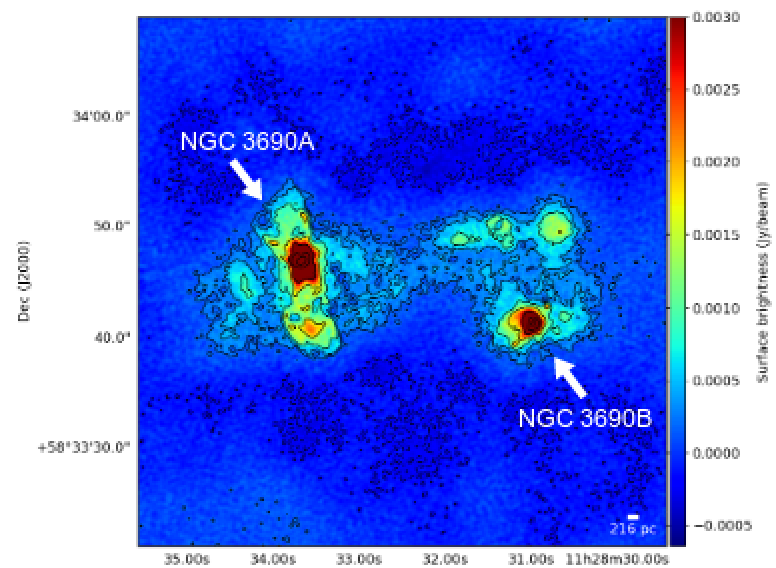
M82 (Varenius et al. 2014) – LOFAR 150MHz



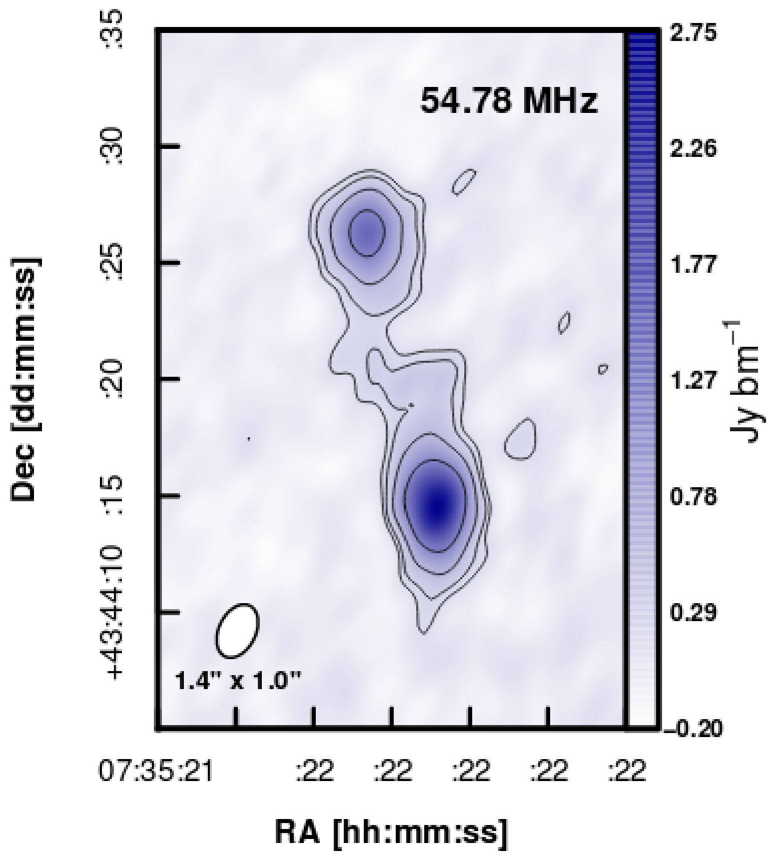
Arp220 (Varenius et al. 2016)

Subarcsecond resolution crucial for SF regions
Extended steep-spectrum emission from outflows visible at low frequencies
Spectral indices – SNRs vs HII regions

Arp299 (Ramirez-Olivencia 2017 in prep)
See poster by Varenius et al.



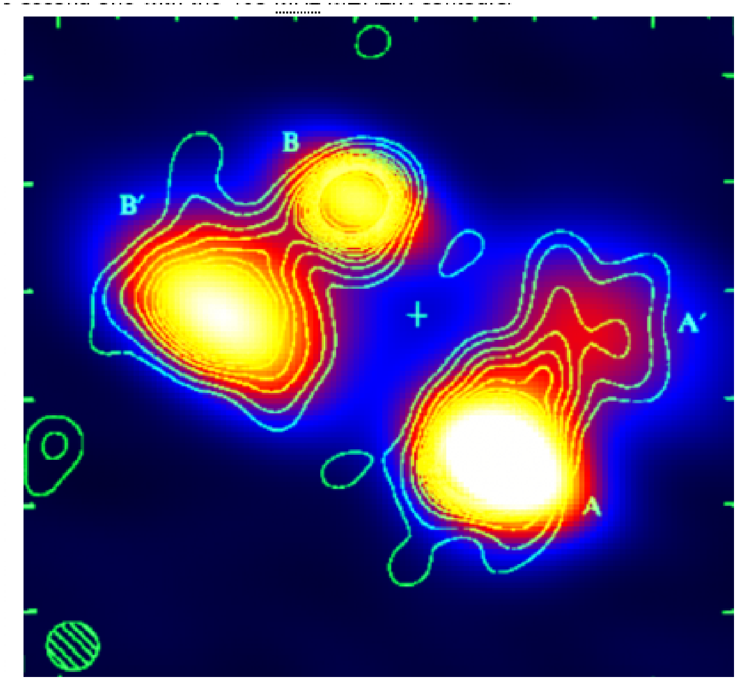
Science with the long baselines: radio galaxies



Morabito et al. 2016 (LBA! 50MHz)

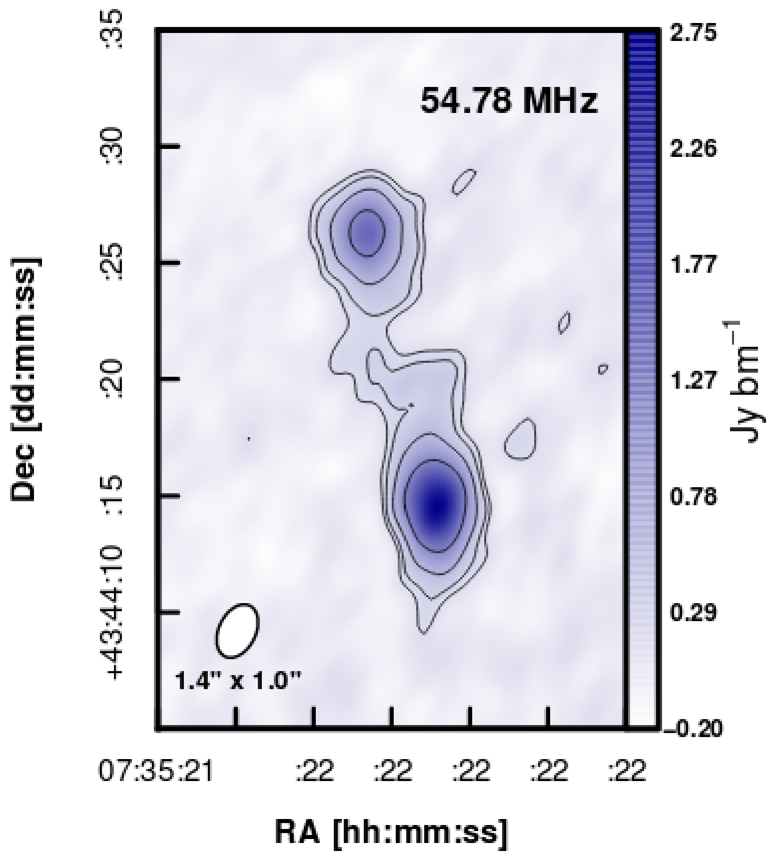
Steep-spectrum emission:

- nature of USS radio galaxies
- spectral indices/physics of extended radio emission



3C196 (Wucknitz 2012)

Science with the long baselines: radio galaxies



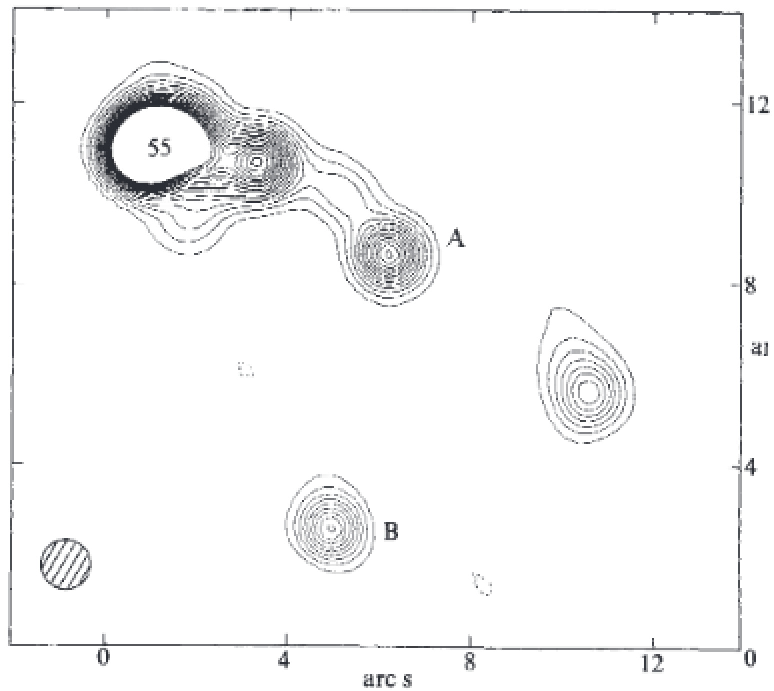
Steep-spectrum emission:

- nature of USS radio galaxies
- spectral indices/physics of extended radio emission

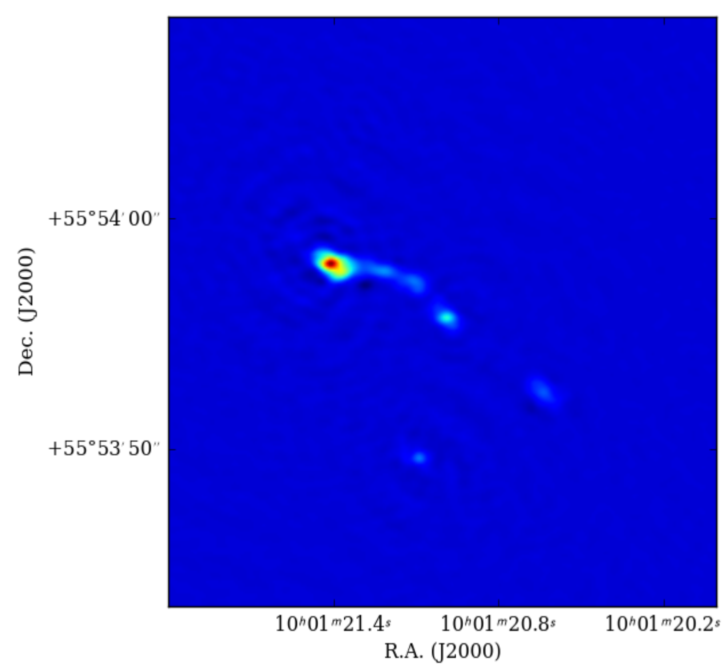
Morabito et al. 2016 (LBA! 50MHz)

Technically very difficult – extreme ionosphere/calibration problems
But seriously unexplored part of parameter space at high resolution

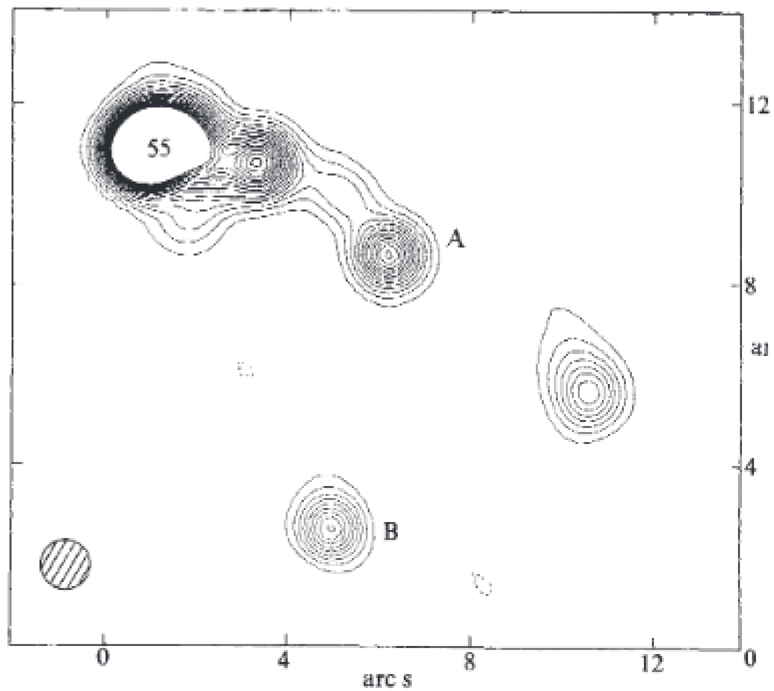
Science with the long baselines: gravitational lenses



Q0957+561 (Noble et al. 1988, 408MHz)
LOFAR LB (Hartley et al. in prep, 151 MHz)



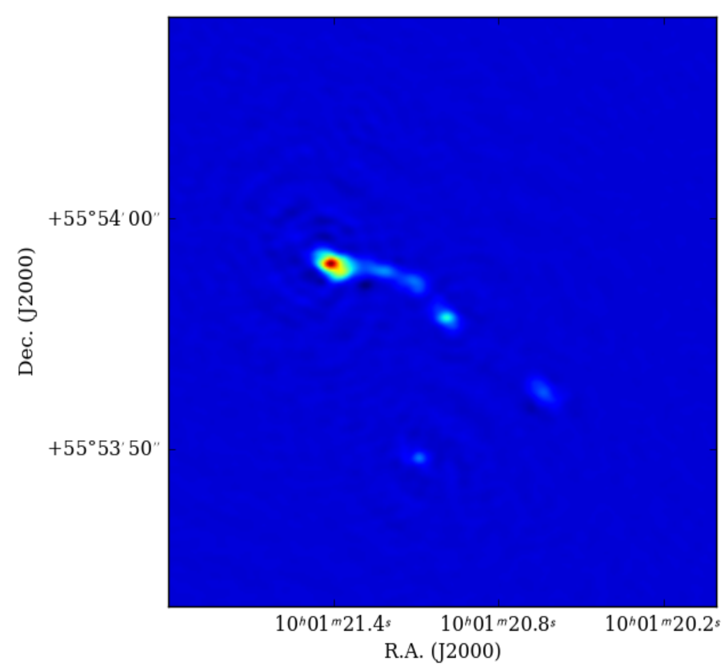
Science with the long baselines: gravitational lenses



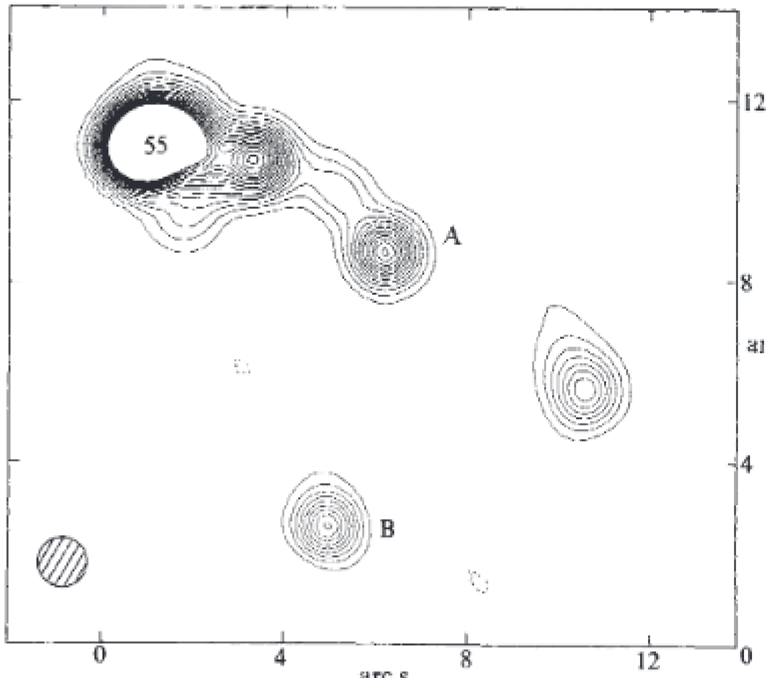
Q0957+561 (Noble et al. 1988, 408MHz)
LOFAR LB (Hartley et al. in prep, 151 MHz)

Science from comparison of 2 sightlines, 1 object

- foreground effects / scattering in lens elliptical
- potential for surveys with rare objects



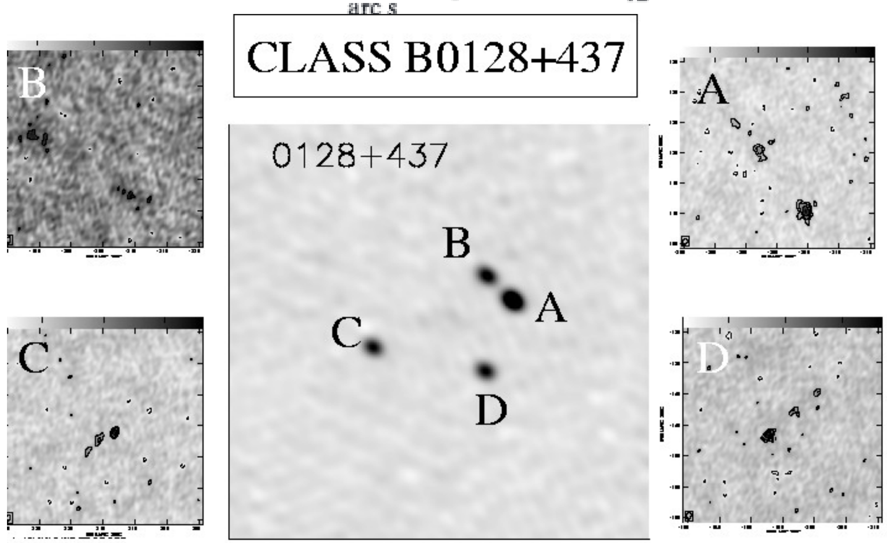
Science with the long baselines: gravitational lenses



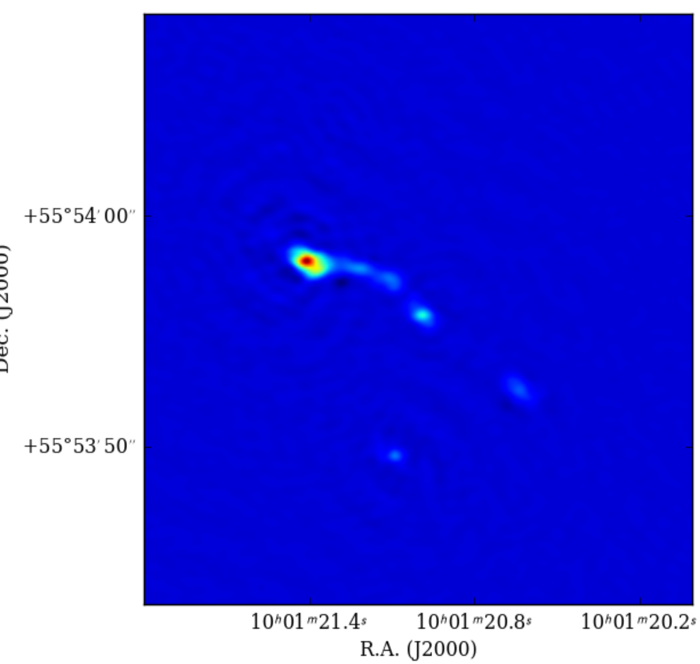
Q0957+561 (Noble et al. 1988, 408MHz)
LOFAR LB (Hartley et al. in prep, 151 MHz)

Science from comparison of 2 sightlines, 1 object

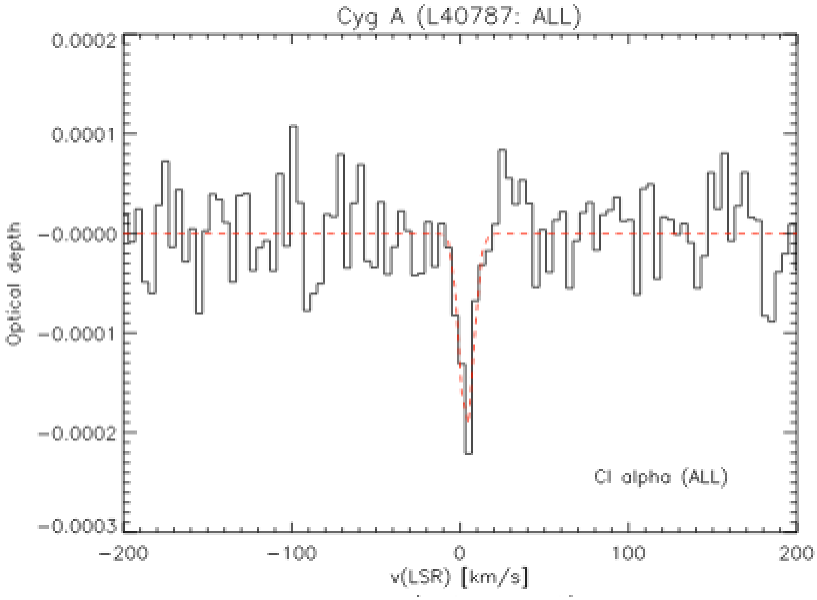
- foreground effects / scattering in lens elliptical
- potential for surveys with rare objects



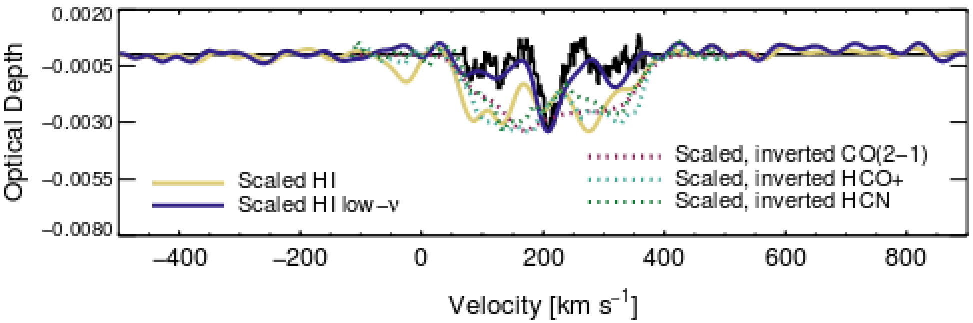
MERLIN full-track+VLBA 2hrs (all at 5GHz)



Science with long baselines: radio recombination lines



CRRL Cyg A (Oonk et al. 2014)



CRRL M82 (Morabito et al. 2014)

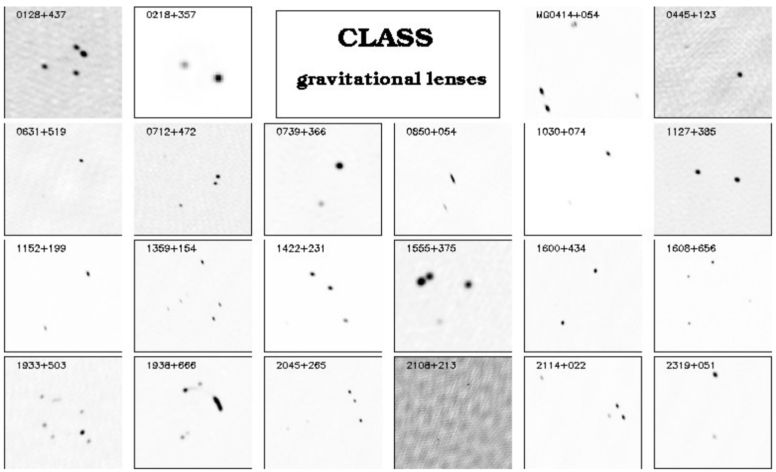
Studies of cold neutral gas → geometry, physics of CNM
Current work with Dutch array
With LB, spatially resolved absorption (cf. current HI)

Important to go from single-object science to large samples

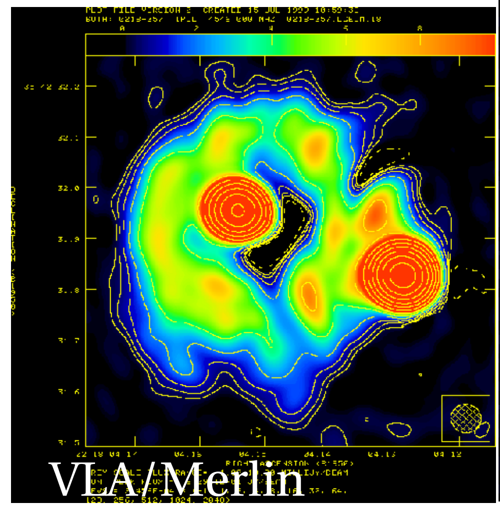
- Discovery of rare objects (e.g. strong lensing surveys)
- Census of AGN at high resolution (census of BH, resolution important for compact sources)
- Important to have large samples for statistics of radio source evolution

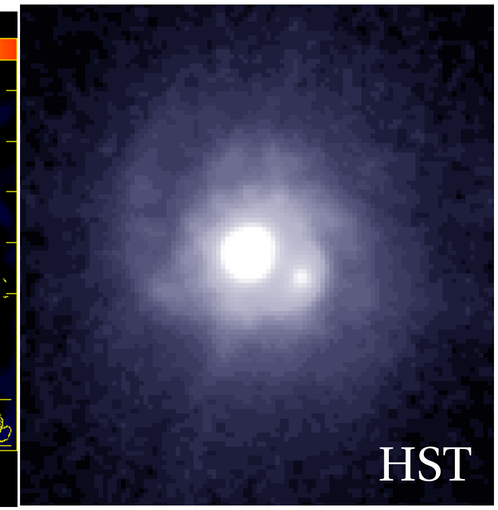
Strong gravitational lens surveys

- Important for studies of dark matter/substructures
- Radio lenses especially useful for substructure with VLBI/central images
- Discovery is difficult – objects are rare

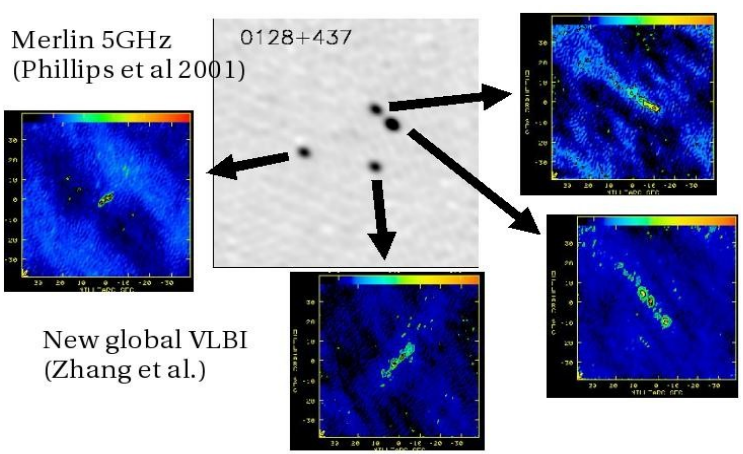


Samples still dominated by 1990's surveys (MG/CLASS – total of about 40)





Science example: CLASS0218+357 (H0 measurement/mass models)



0128+437 (substructure/foregrounds)

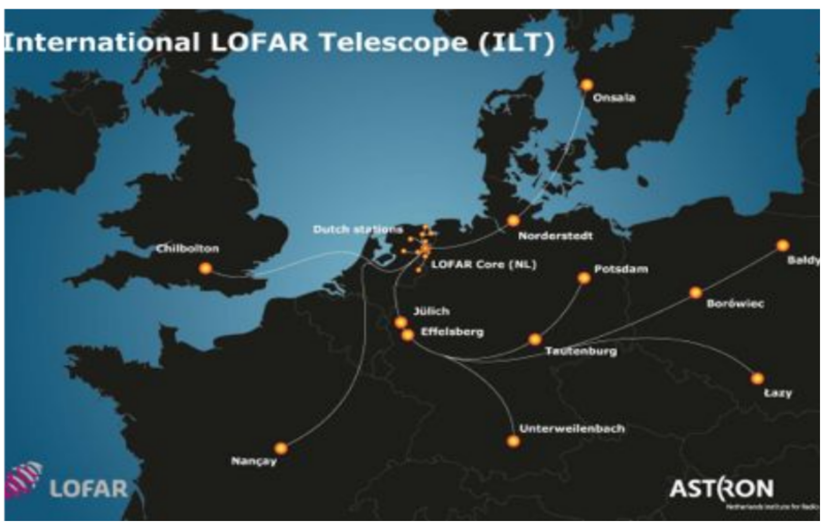
3. Towards LOFAR-LB large scale imaging:

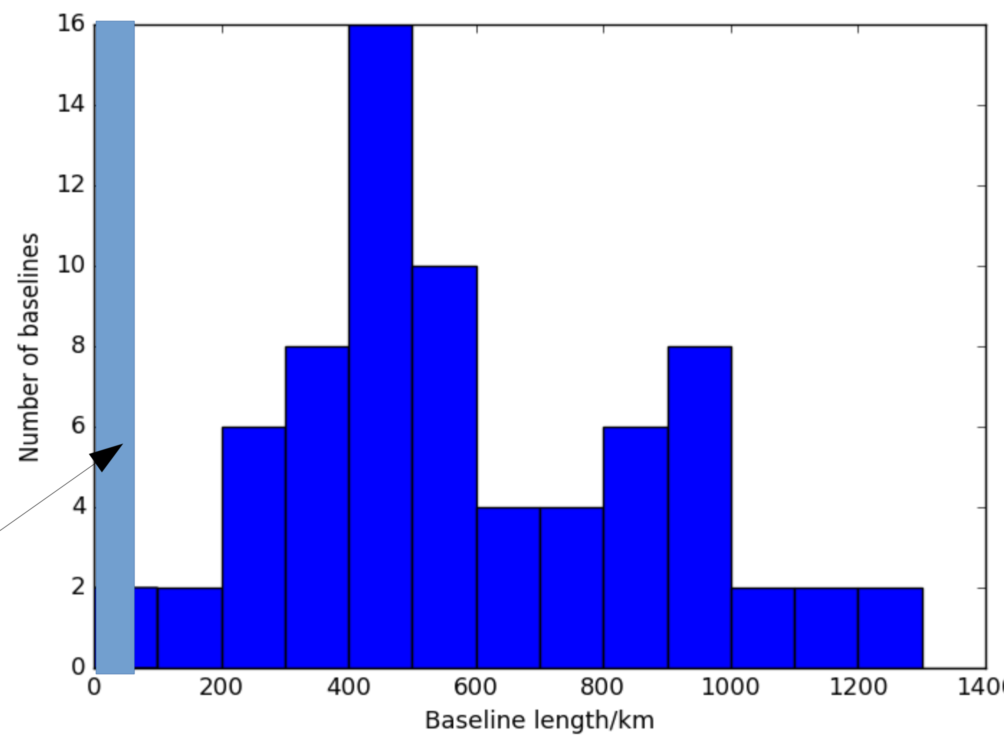
The LBCS calibrator survey

Need calibrator sources for delay/phase cal, before imaging

Needs to be dense net (ionospheric coherence patch ~1 deg or less)
Needs to contain sources with bright compact structure (hard at 150MHz)

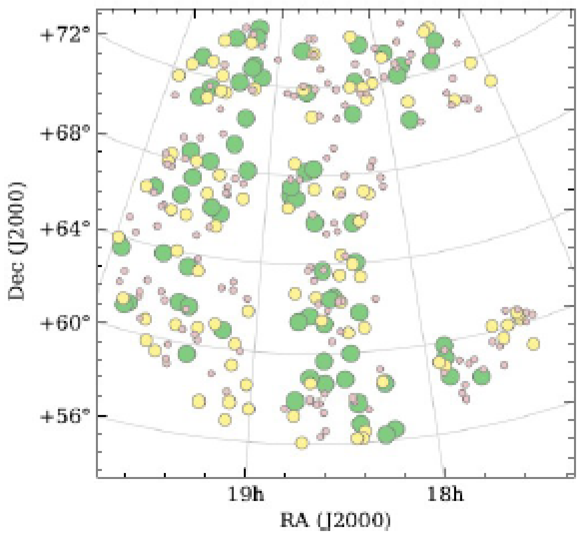
PL/IR baselines

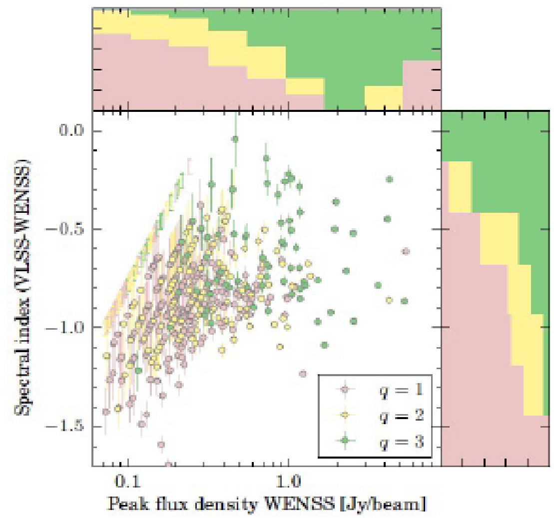




NL baselines

Pilot survey (Moldon et al. 2014)

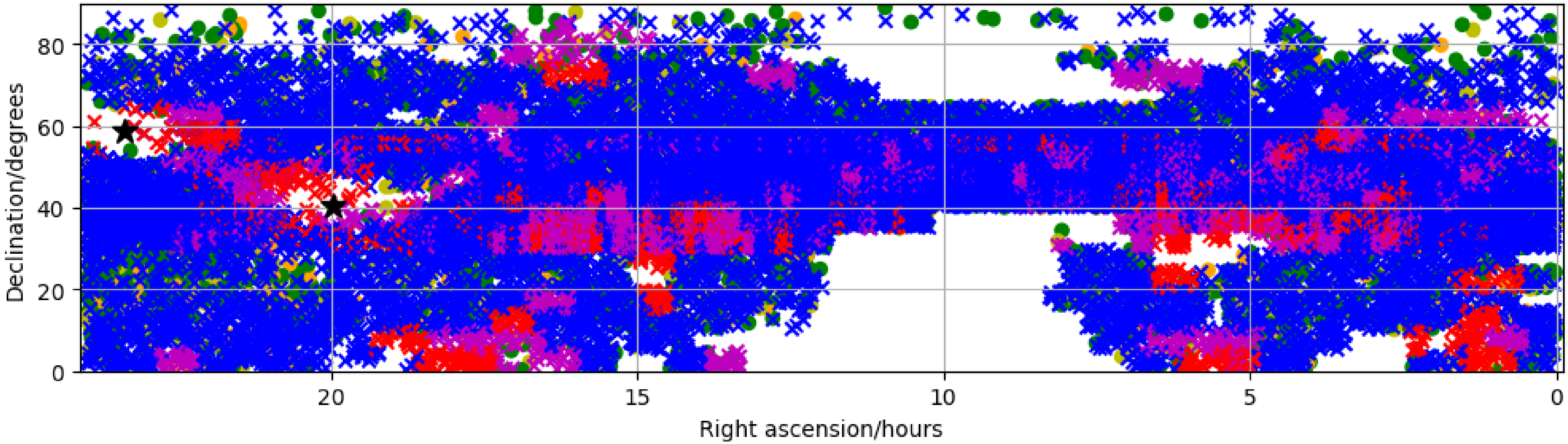




Snapshot (3-min) observation, 30 sources at once
No imaging, but just see if fringe-fit to a point source works (green) or not (red)
Brighter sources are better
Flat-spectrum sources are better
20-40% success rate
Extended to LBCS (all northern sky), 96h observations, now complete
Data reduction underway

Technical challenges – calibration

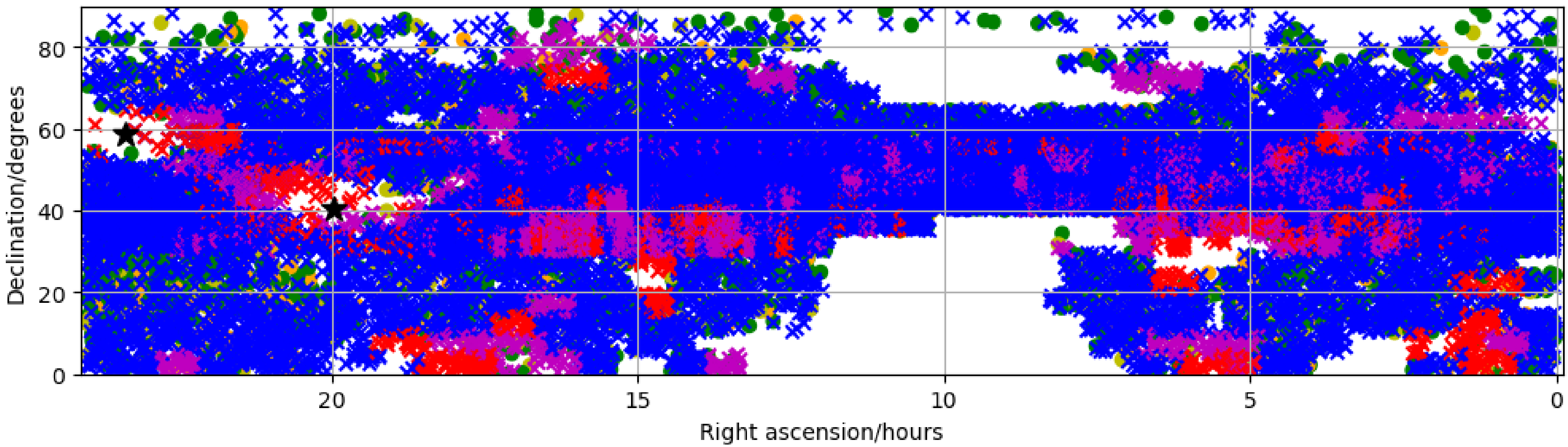
- Calibrator sources necessary for ionospheric phase calibration
- More compact sources for long-baseline calibration are rarer
- LBCS survey under way (LBWG/Jackson et al. 2016)



Current status of LBCS calibrator survey (early 2018)

Technical challenges – calibration

- Calibrator sources necessary for ionospheric phase calibration
- More compact sources for long-baseline calibration are rarer
- LBCS survey under way (LBWG/Jackson et al. 2016)

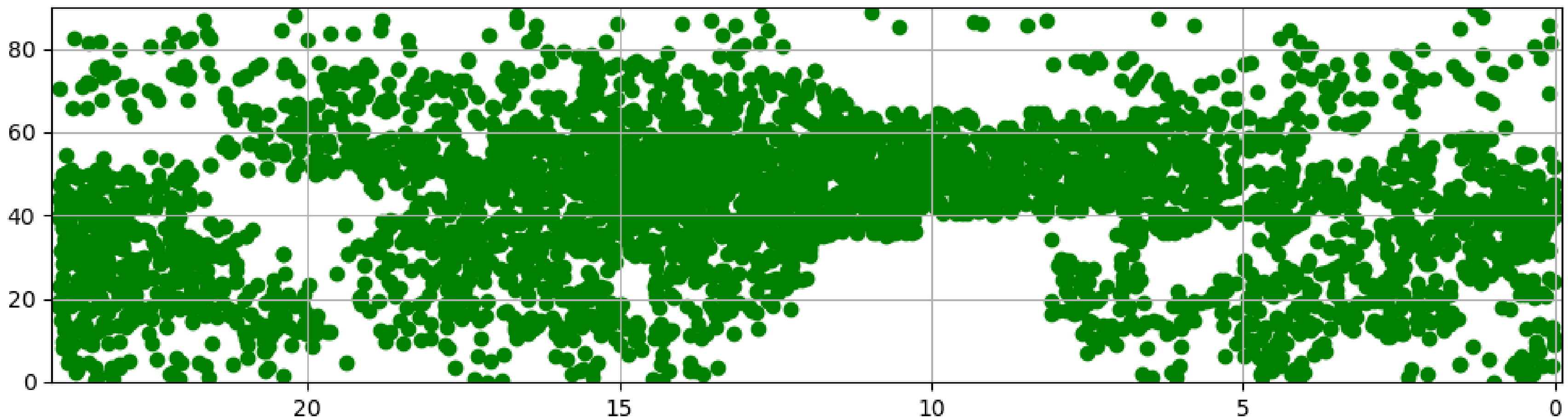


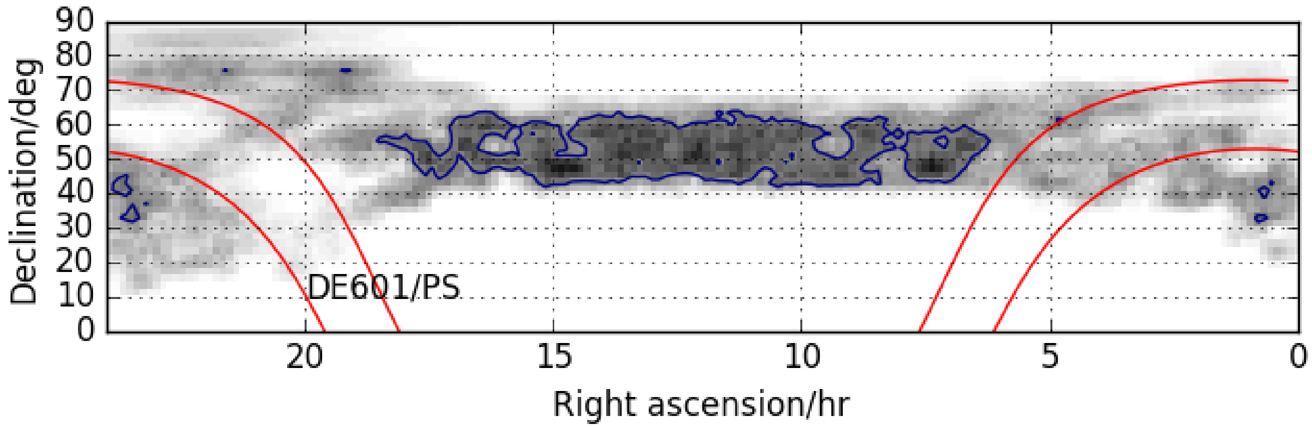
Current status of LBCS calibrator survey (early 2018)

Calibrator list available on vo.astron.nl
Updated as-it-happens version on http://www.jb.man.ac.uk/~njj/lbcs_stats.sum

Technical challenges – calibration

- Calibrator sources necessary for ionospheric phase calibration
- More compact sources for long-baseline calibration are rarer
- LBCS survey under way (LBWG/Jackson et al. 2016)

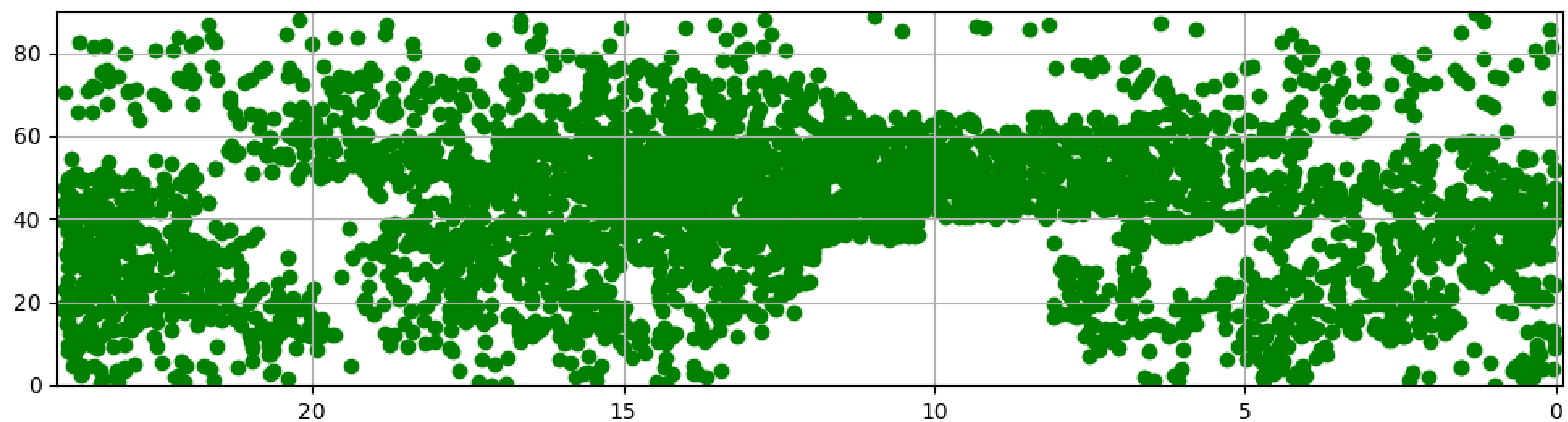


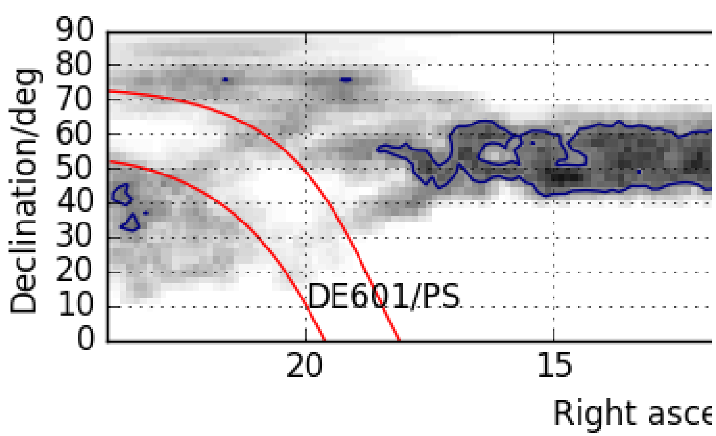


200-km baselines
1 source/sq deg

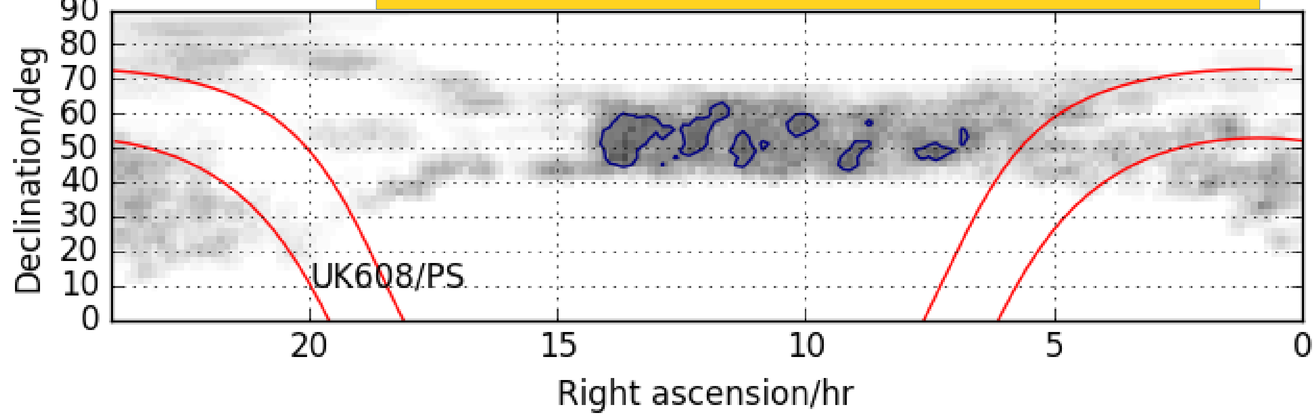
Technical challenges – calibration

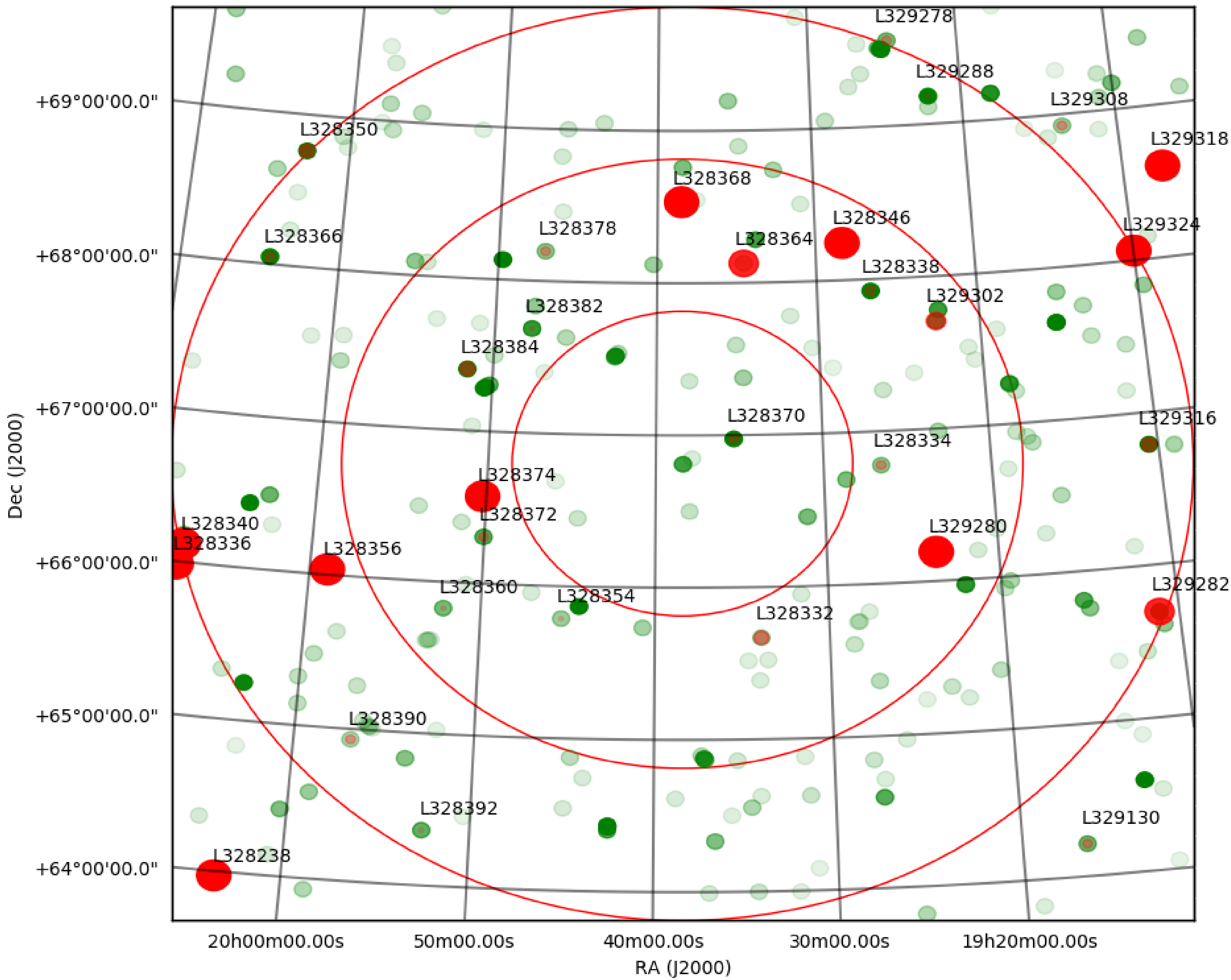
- Calibrator sources necessary for ionospheric phase calibration
- More compact sources for long-baseline calibration are rarer
- LBCS survey under way (LBWG/Jackson et al. 2016)





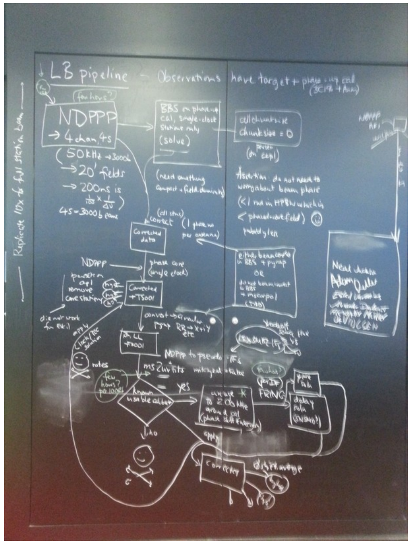
800-km baselines - 0.5 source/sq deg



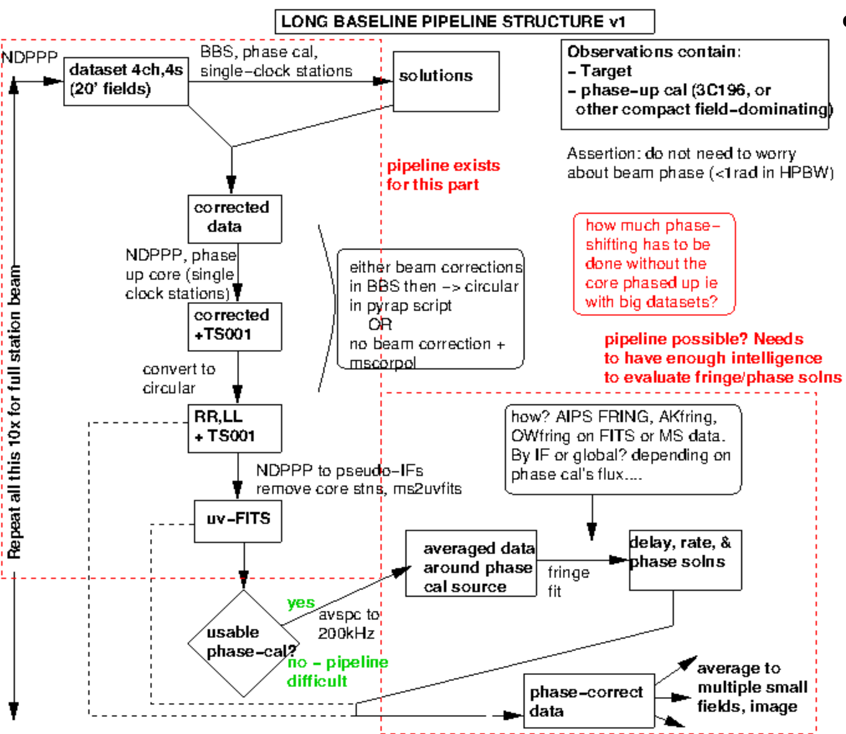

L329278
L329288
L329308
L329318
L328350
L328368
L328346
L329324
L328366
L328378
L328364
L328338
L329302
L328382
L328384
L328370
L329316
L328334
L328374
L328372
L328340
L328336
L328356
L329280
L328360
L328354
L328332
L329282
L328390
L328392
L329130
L328238
+69°00'00.0"
+68°00'00.0"
+67°00'00.0"
+66°00'00.0"
+65°00'00.0"
+64°00'00.0"
Dec (J2000)
20h00m00.00s
50m00.00s
40m00.00s
30m00.00s
19h20m00.00s
RA (J2000)

4. Towards large-scale imaging with LOFAR LB

see also talks by Marco Iacobelli (LB pipeline)
and Leah Morabito (LoTSS pipeline processing)



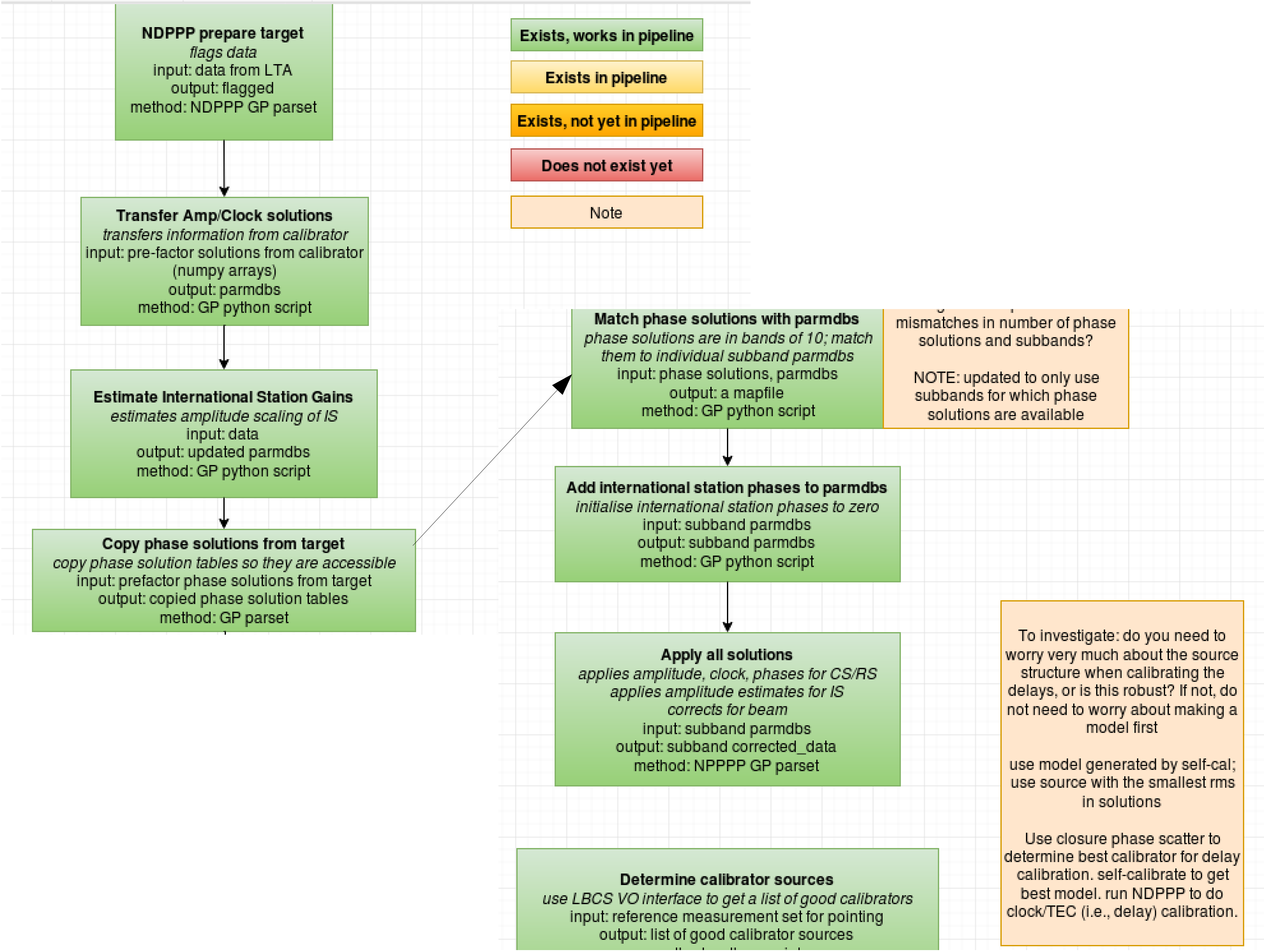
Current technical capabilities: pipeline processing

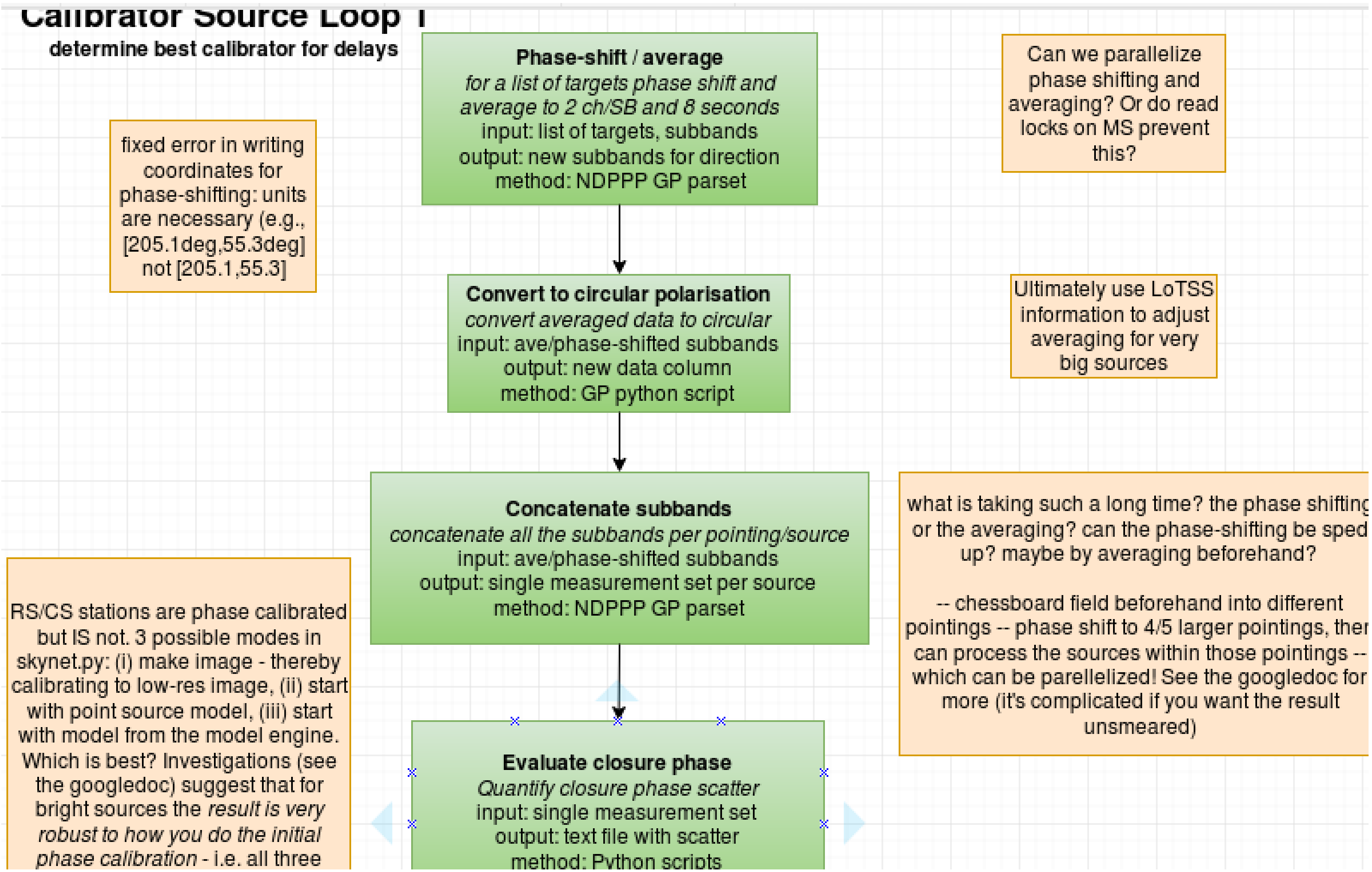


- Pre-processing – shared with surveys
- LBWG pipeline – superterp formation, circular conversio
- Delay calibration – manual, being pipelined
- Shift/average to sources – being pipelined
- Phase calibration/imaging – being pipelined

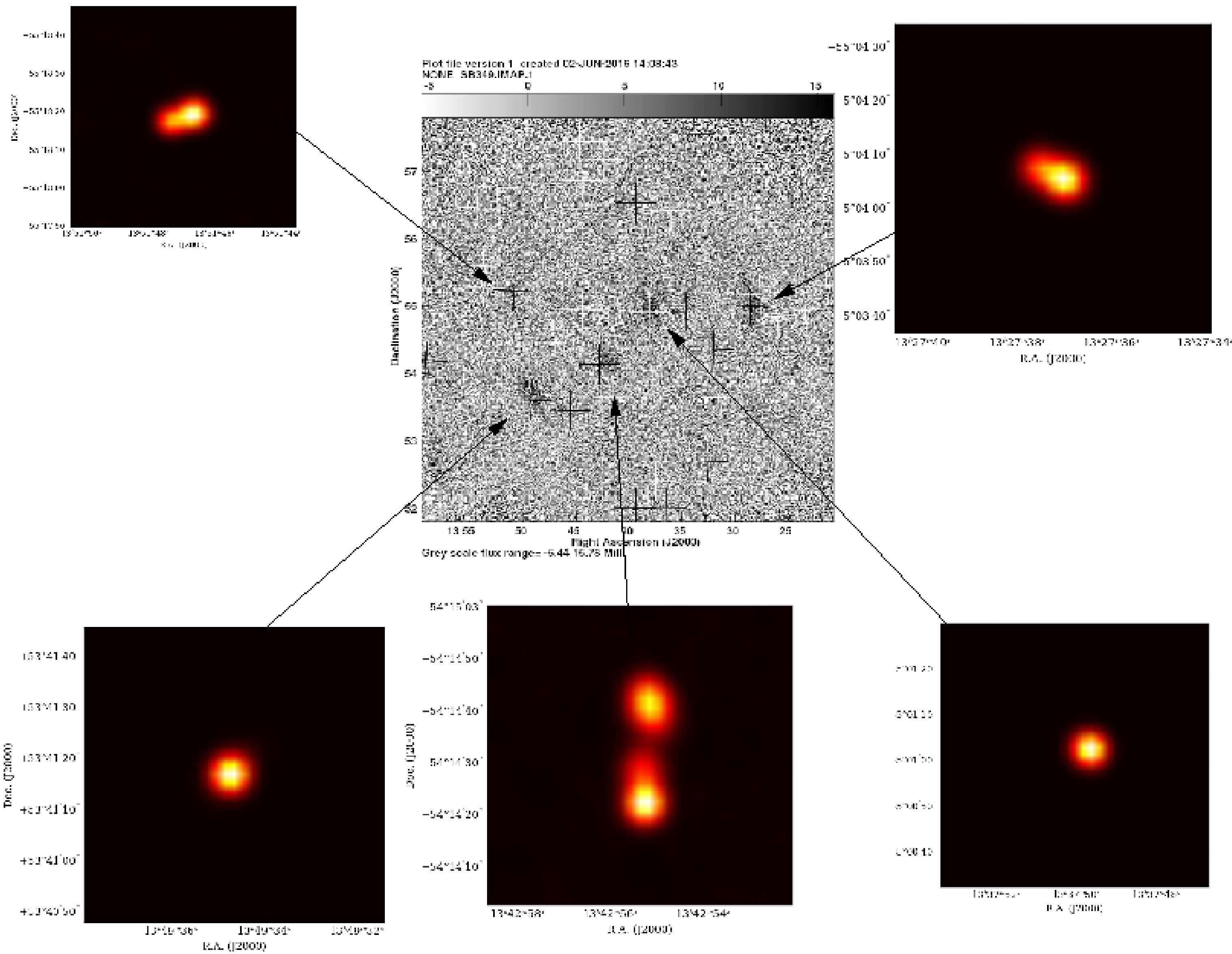
Limitations

- Storage, processing power (IB datasets 50Gb/sb 16ch/1s
- Ionospheric calibration
- Unclear how faint we can go
- Close to thermal noise demonstrated however

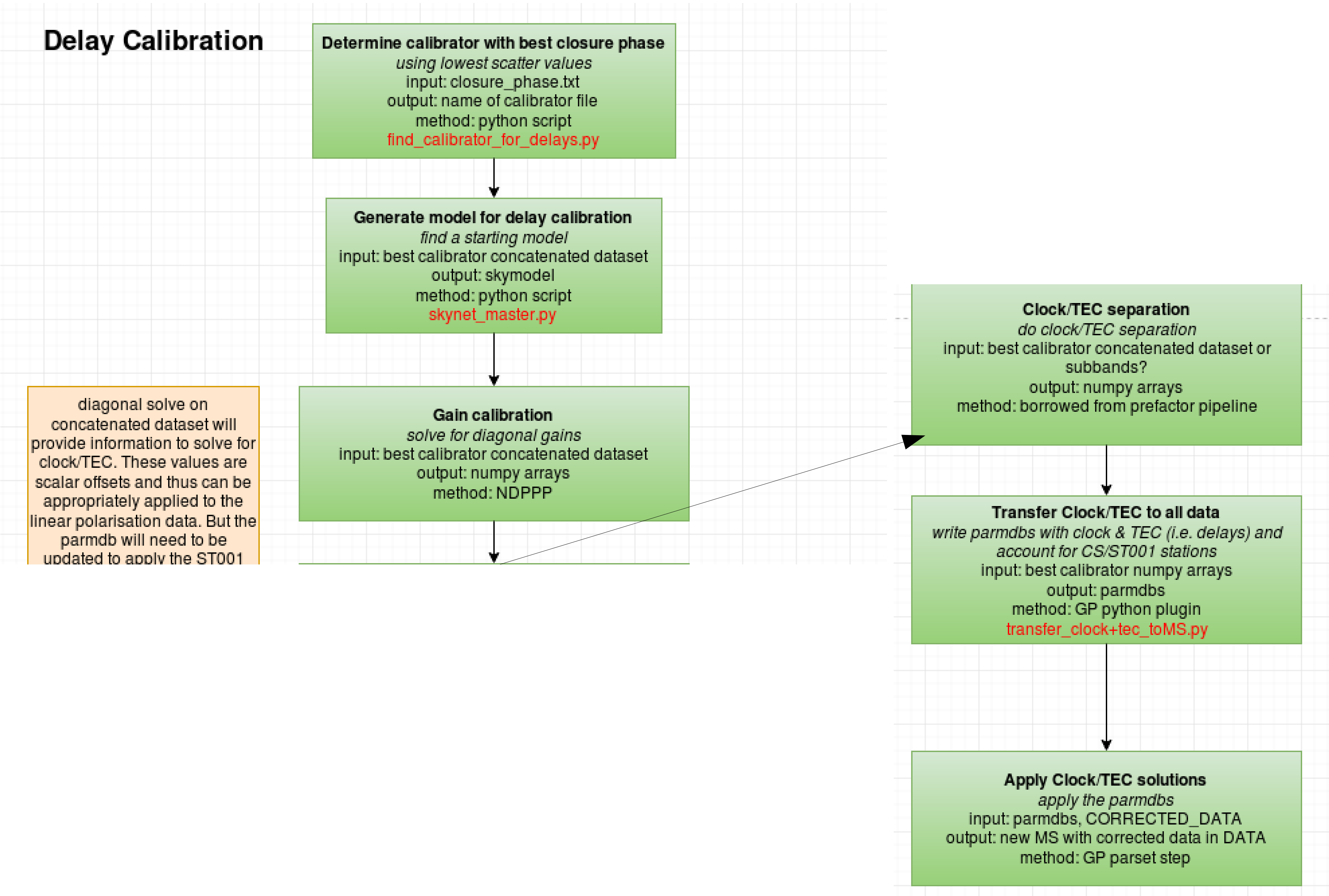

NDPPP prepare target
flags data
input: data from LTA
output: flagged
method: NDPPP GP parset
Exists, works in pipeline
Exists in pipeline
Exists, not yet in pipeline
Does not exist yet
Note
Transfer Amp/Clock solutions
transfers information from calibrator
input: pre-factor solutions from calibrator (numpy arrays)
output: parmdbs
method: GP python script
Estimate International Station Gains
estimates amplitude scaling of IS
input: data
output: updated parmdbs
method: GP python script
Copy phase solutions from target
copy phase solution tables so they are accessible
input: prefactor phase solutions from target
output: copied phase solution tables
method: GP parset
Match phase solutions with parmdbs
phase solutions are in bands of 10; match them to individual subband parmdbs
input: phase solutions, parmdbs
output: a mapfile
method: GP python script
mismatches in number of phase solutions and subbands?
NOTE: updated to only use subbands for which phase solutions are available
Add international station phases to parmdbs
initialise international station phases to zero
input: subband parmdbs
output: subband parmdbs
method: GP python script
Apply all solutions
applies amplitude, clock, phases for CS/RS
applies amplitude estimates for IS
corrects for beam
input: subband parmdbs
output: subband corrected_data
method: NPPPP GP parset
To investigate: do you need to worry very much about the source structure when calibrating the delays, or is this robust? If not, do not need to worry about making a model first
use model generated by self-cal; use source with the smallest rms in solutions
Use closure phase scatter to determine best calibrator for delay calibration. self-calibrate to get best model. run NDPPP to do clock/TEC (i.e., delay) calibration.
Determine calibrator sources
use LBCS VO interface to get a list of good calibrators
input: reference measurement set for pointing
output: list of good calibrator sources

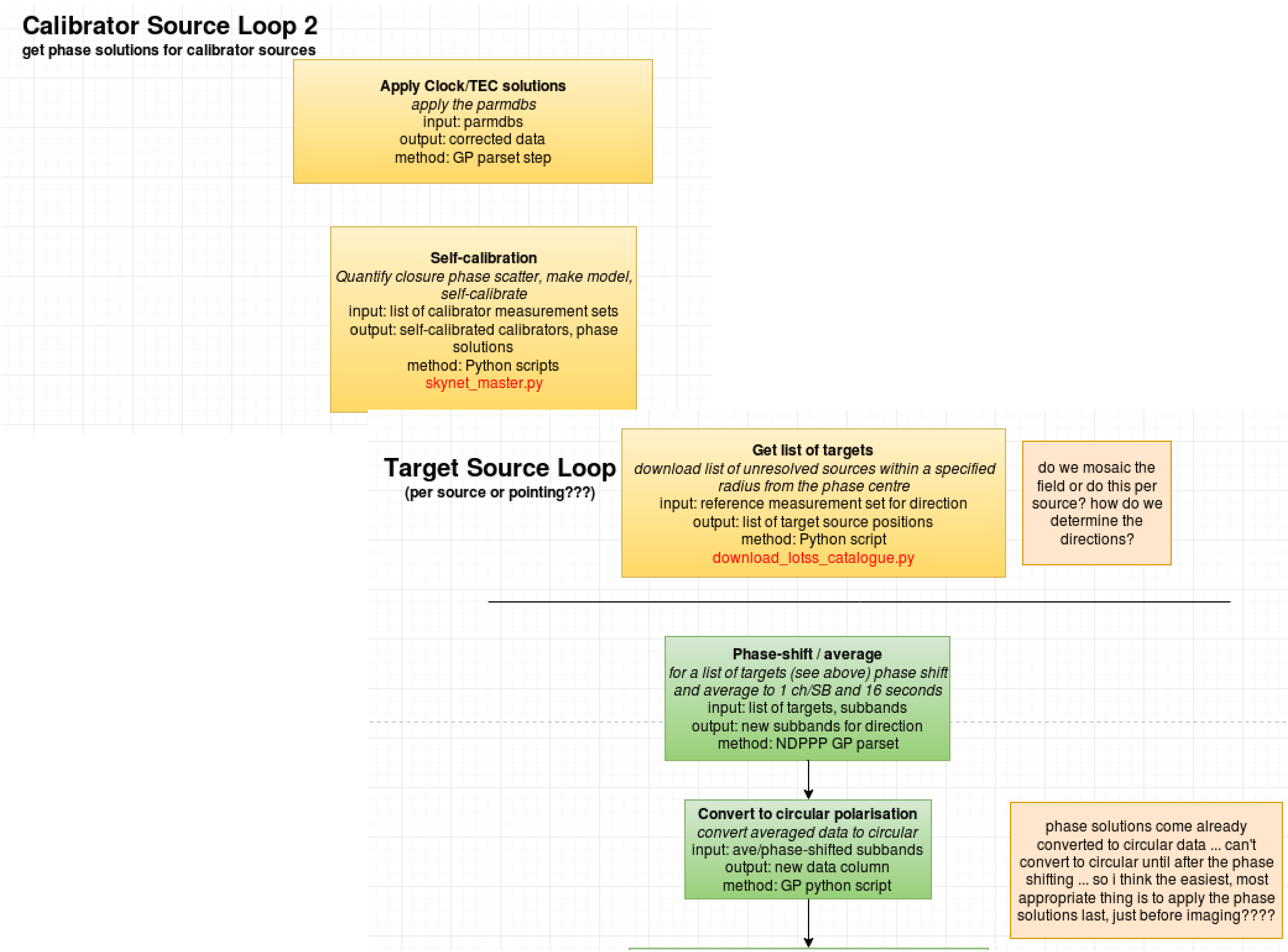

Calibrator Source Loop 1
determine best calibrator for delays
Phase-shift / average
for a list of targets phase shift and average to 2 ch/SB and 8 seconds
input: list of targets, subbands
output: new subbands for direction
method: NDPPP GP parset
Can we parallelize phase shifting and averaging? Or do read locks on MS prevent this?
fixed error in writing coordinates for phase-shifting: units are necessary (e.g., [205.1deg,55.3deg] not [205.1,55.3]
Convert to circular polarisation
convert averaged data to circular
input: ave/phase-shifted subbands
output: new data column
method: GP python script
Ultimately use LoTSS information to adjust averaging for very big sources
Concatenate subbands
concatenate all the subbands per pointing/source
input: ave/phase-shifted subbands
output: single measurement set per source
method: NDPPP GP parset
what is taking such a long time? the phase shifting or the averaging? can the phase-shifting be sped up? maybe by averaging beforehand?
-- chessboard field beforehand into different pointings -- phase shift to 4/5 larger pointings, then can process the sources within those pointings -- which can be parellelized! See the googledoc for more (it's complicated if you want the result unsmeared)
RS/CS stations are phase calibrated but IS not. 3 possible modes in skynet.py: (i) make image - thereby calibrating to low-res image, (ii) start with point source model, (iii) start with model from the model engine. Which is best? Investigations (see the googledoc) suggest that for bright sources the result is very robust to how you do the initial phase calibration - i.e. all three
Evaluate closure phase
Quantify closure phase scatter
input: single measurement set
output: text file with scatter
method: Python scripts

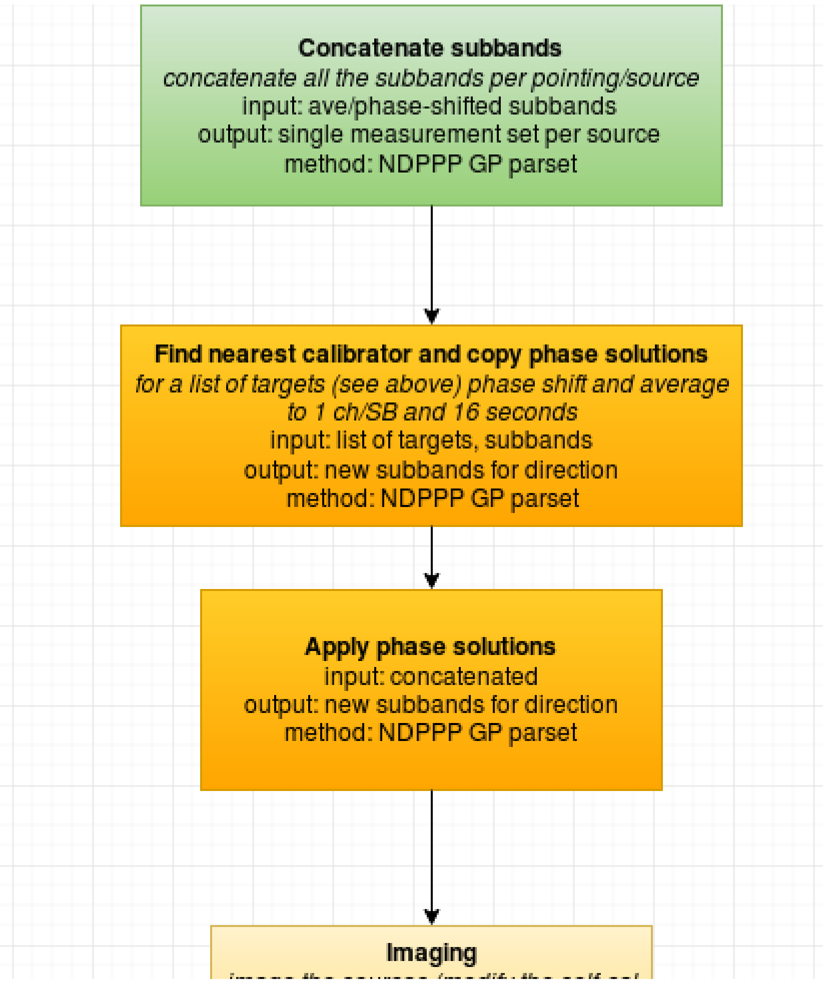
Current technical capabilities – illustration on LBWG surveys field



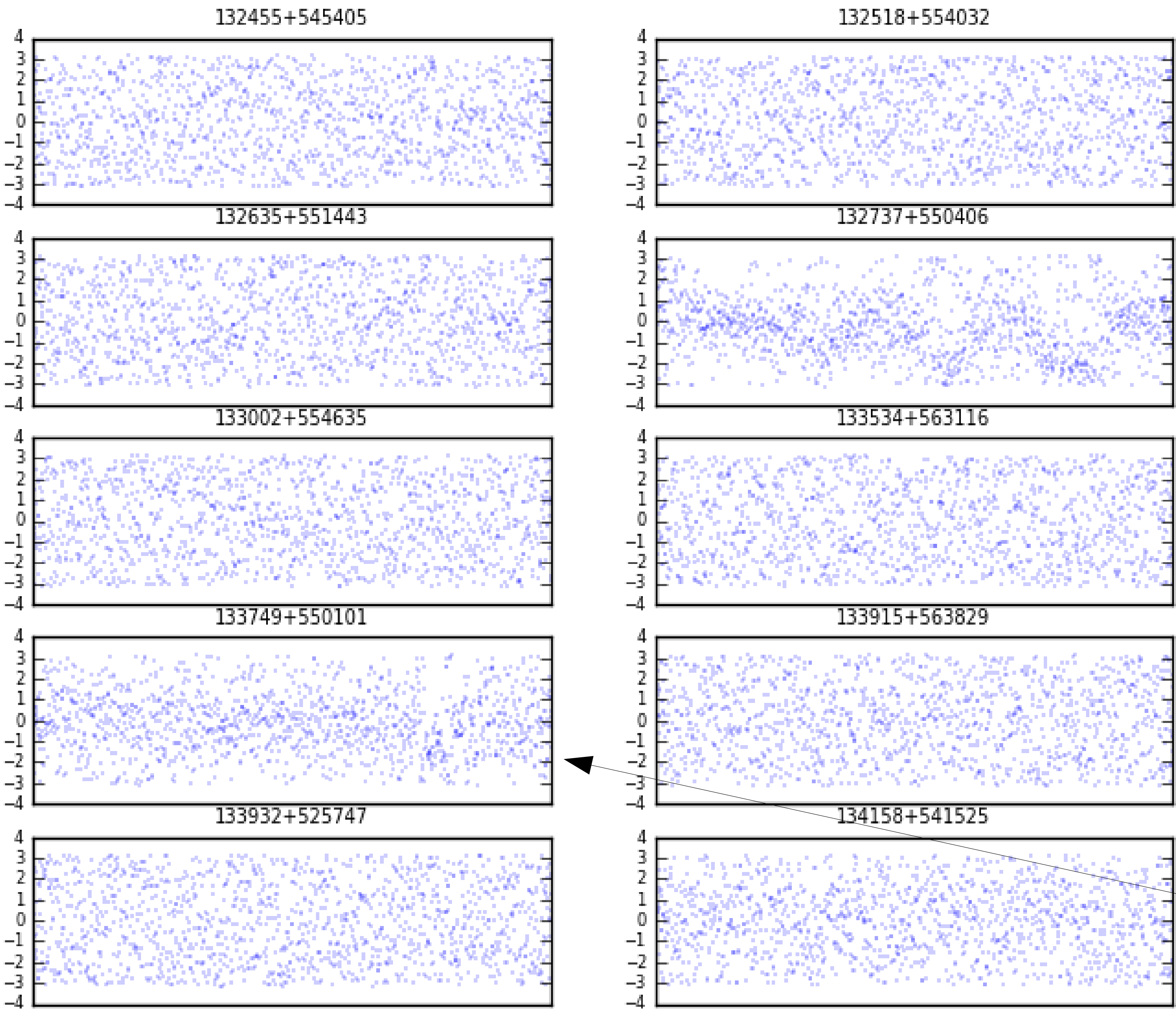
Greyscale: fringe-rate/delay map on ST001 – DE605 baseline
Cutouts: FIRST images of some of the brighter sources in the field


Delay Calibration
Determine calibrator with best closure phase
using lowest scatter values
input: closure_phase.txt
output: name of calibrator file
method: python script
find_calibrator_for_delays.py
Generate model for delay calibration
find a starting model
input: best calibrator concatenated dataset
output: skymodel
method: python script
skynet_master.py
diagonal solve on concatenated dataset will provide information to solve for clock/TEC. These values are scalar offsets and thus can be appropriately applied to the linear polarisation data. But the parmdb will need to be updated to apply the ST001
Gain calibration
solve for diagonal gains
input: best calibrator concatenated dataset
output: numpy arrays
method: NDPPP
Clock/TEC separation
do clock/TEC separation
input: best calibrator concatenated dataset or subbands?
output: numpy arrays
method: borrowed from prefactor pipeline
Transfer Clock/TEC to all data
write parmdbs with clock & TEC (i.e. delays) and account for CS/ST001 stations
input: best calibrator numpy arrays
output: parmdbs
method: GP python plugin
transfer_clock+tec_toMS.py
Apply Clock/TEC solutions
apply the parmdbs
input: parmdbs, CORRECTED_DATA
output: new MS with corrected data in DATA
method: GP parset step


Calibrator Source Loop 2
get phase solutions for calibrator sources
Apply Clock/TEC solutions
apply the parmdbs
input: parmdbs
output: corrected data
method: GP parset step
Self-calibration
Quantify closure phase scatter, make model, self-calibrate
input: list of calibrator measurement sets
output: self-calibrated calibrators, phase solutions
method: Python scripts
skynet_master.py
Target Source Loop
(per source or pointing???)
Get list of targets
download list of unresolved sources within a specified radius from the phase centre
input: reference measurement set for direction
output: list of target source positions
method: Python script
download_lotss_catalogue.py
do we mosaic the field or do this per source? how do we determine the directions?
Phase-shift / average
for a list of targets (see above) phase shift and average to 1 ch/SB and 16 seconds
input: list of targets, subbands
output: new subbands for direction
method: NDPPP GP parset
Convert to circular polarisation
convert averaged data to circular
input: ave/phase-shifted subbands
output: new data column
method: GP python script
phase solutions come already converted to circular data ... can't convert to circular until after the phase shifting ... so i think the easiest, most appropriate thing is to apply the phase solutions last, just before imaging????


Concatenate subbands
concatenate all the subbands per pointing/source
input: ave/phase-shifted subbands
output: single measurement set per source
method: NDPPP GP parset
Find nearest calibrator and copy phase solutions
for a list of targets (see above) phase shift and average to 1 ch/SB and 16 seconds
input: list of targets, subbands
output: new subbands for direction
method: NDPPP GP parset
Apply phase solutions
input: concatenated
output: new subbands for direction
method: NDPPP GP parset
Imaging

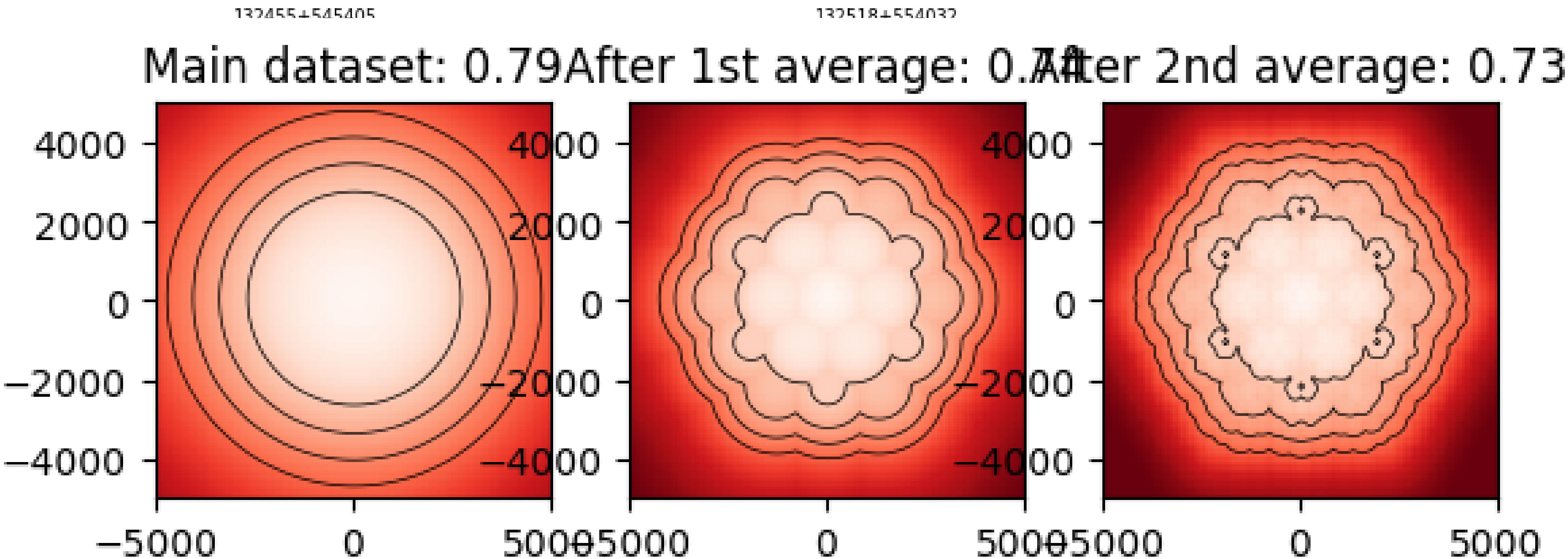
After delay cal, large dataset is then shifted and averaged around particular sources
Then form the closure phase.



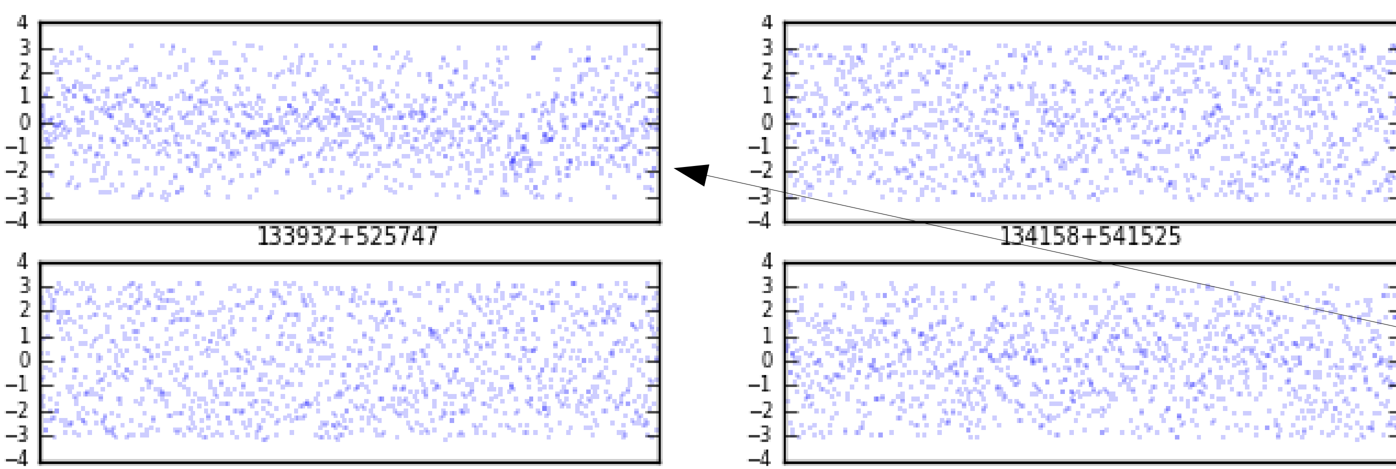
Used this one to start!
(at western edge of field)

This one could have been
used if necessary

Large dataset is then shifted and averaged around particular sources
(if many sources, need to do chessboard averaging to avoid huge runtimes and smearing)

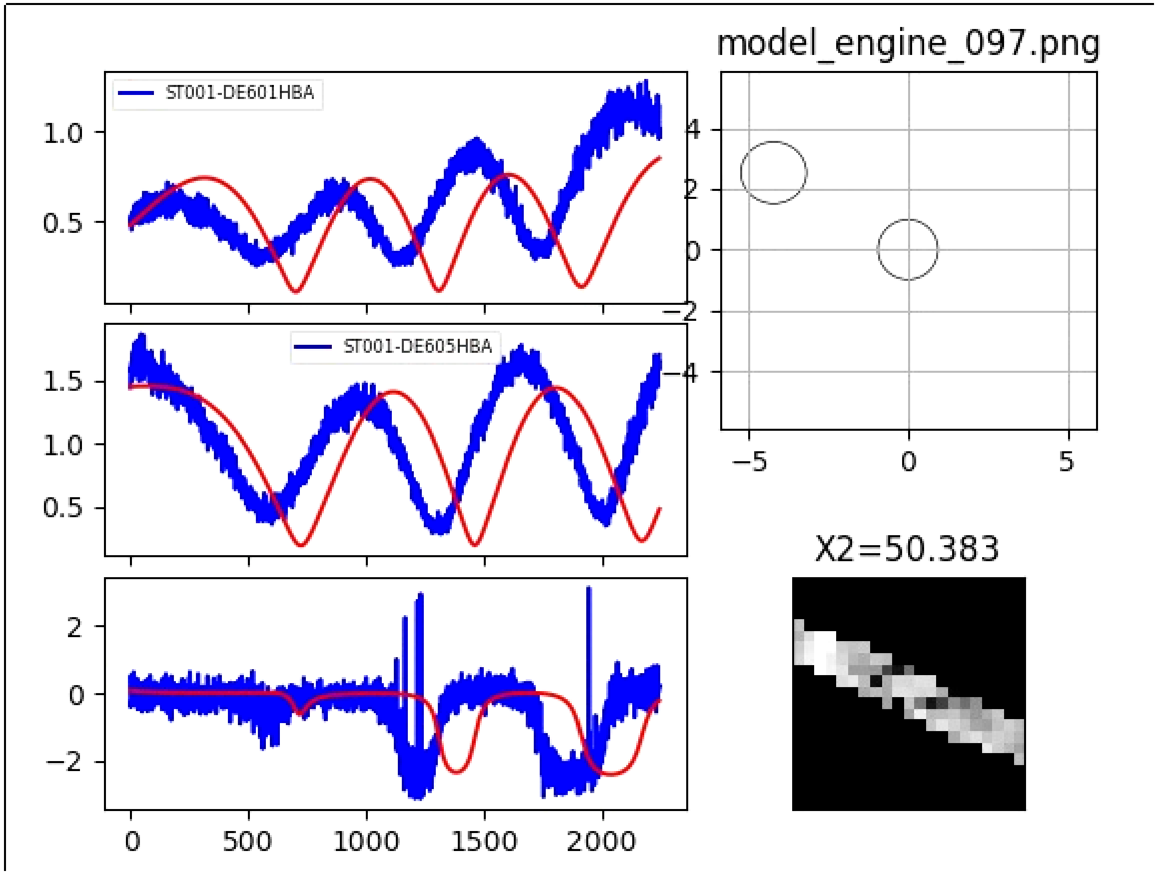


d)



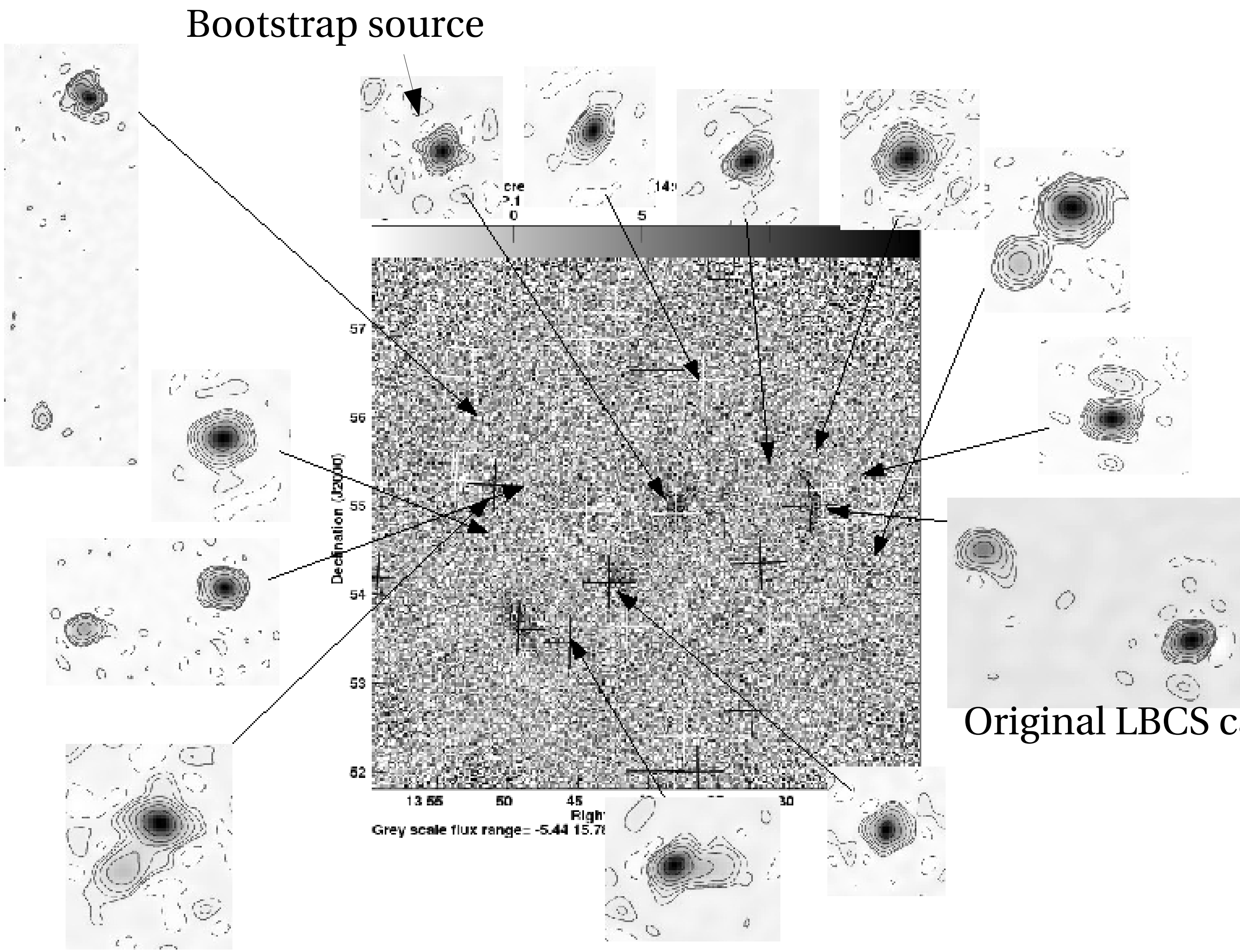
This one could have been used if necessary

Starting models can be made using closure phases



Bright sources relatively robust against starting model
Faint sources are not – major issue in avoiding selfcaling sources into existence
- here important to examine closure phases

Then propagate phase solutions across field from initial calibrator to bootstrap sources



Are we nearly there yet?

- LBWG working on pipeline (see Marco talk after lunch)
- Preprocessing and delay calibration under control
- Currently writing final loop for going around the field imaging/selfcal